AN INVESTIGATION OF THE PROPERTIES OF THE EXPONENTIAL MOVING AVERAGE POINT PROCESS

Tzy-dah Jathro Lo

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

AN INVESTIGATION OF THE PROPERTIES
OF THE
EXPONENTIAL MOVING AVERAGE POINT PROCESS

by

Lo, Tzy-dah Jathro

March 1976

Thesis Advisor: P. A. W. Lewis

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Investigation of the Properties of the Exponential Moving Average Point Process		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis March 1976
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Lo, Tzy-dah Jathro		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1976
		13. NUMBER OF PAGES 57
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Linear Combinations
Moving Average
Point Process
Stationary Random Sequence
Estimating the Parameter β
Serial Correlations
Joint Expectations
Asymptotically Unbiased

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Properties of a stationary sequence of random variables $\{X_i\}$ which have exponential marginal distributions and random linear combinations of order one of an i.i.d. exponential sequence $\{\varepsilon_i\}$ were discussed by Lawrence and Lewis (1976); they called this model the EMA1 (exponential moving average of order one) point process. This paper will investigate the estimators of the parameter β of the EMA1 process, and some basic properties of the EMA2 process, and then extend these results to the EMAk process.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601

An Investigation of the Properties
of the
Exponential Moving Average Point Process

by

Lo, Tzy-dah Jathro
Commander, Chinese Navy
B.S., Chinese Naval Academy, 1959
M.S., University of Iowa, U.S., 1971

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1976

ABSTRACT

Properties of a stationary sequence of random variables $\{X_i\}$ which have exponential marginal distributions and random linear combinations of order one of an i.i.d. exponential sequence $\{\varepsilon_i\}$ were discussed by Lawrance and Lewis (1976); they called this model the EMA1 (exponential moving average of order one) point process. This paper will investigate the estimators of the parameter β of the EMA1 process, and some basic properties of the EMA2 process, and then extend these results to the EMAk process.

TABLE OF CONTENTS

LIST OF SYMBOLS

X_i	The *ith* element of the sequence of the time intervals of the point process.
ε_i	The *ith* element of the i.i.d. exponential sequence with parameter λ .
EMAk	Exponential Moving Average of order k.
β_i	Probabilities, $0 \le \beta_i \le 1$, $i=1,2,3,\ldots\ldots$
ρ_j	The *jth* order serial correlations.
T_r	Sum of X_i's; $T_r = X_1 + X_2 + \ldots\ldots + X_r$.
$\phi_r(s)$	Laplace transform of the p.d.f. of T_r; $\phi_r(s) = E(e^{-sT_r})$
p.d.f.	Probability density function.
$F_r(t)$	Distribution function of T_r.
$N_t^{(f)}$	The number of events occurring in the time interval $(0,t]$.
$\psi_f(z;t)$	The generating function of $N_t^{(f)}$; $E(z^{N_t^{(f)}})$.
$\psi_f^*(z;s)$	Laplace transform of $\psi_f(z;t)$.
$m_f^*(t)$	The intensity function of the point process.
$m_f^*(s)$	Laplace transform of $m(t)$.
λ	Parameter of exponential distributions.
s	Variable of Laplace transform.
$f^*_{X_i}(s)$	Laplace transform of the p.d.f. of X_i; $E(e^{-sX_i})$.
$f^{**}_{X_i,X_{i+1}}(s_1,s_2)$	Laplace transform of the joint p.d.f. of X_i and X_{i+1}.
$\hat{\beta}$, $\hat{\hat{\beta}}$	Estimators of β .

Y'	Random variable Y divided by its mean value, i.e. $Y'=Y/\mu y$.
$\hat{\rho}_j$	Estimator of ρ_j .
$\psi(s)$	Laplace transform of the p.d.f. of the ε_i distribution.
$f_{\beta_1\varepsilon_i,\beta_1\varepsilon_{i+1}}$	Joint p.d.f. of $\beta_1\varepsilon_i$ and $\beta_1\varepsilon_{i+1}$.
$X_i^{(k)}$	X_i of the EMAk process.
$\rho_j^{(k)}$	The <u>jth</u> order serial correlation of the EMAk process.

ACKNOWLEDGEMENT

I am deeply indebted to Professor P. A. W. Lewis for his invaluable assistance and numerous helpful directions throughout the entire period that I was working on this thesis. I must thank Lieutenant Commander M. G. Bell, Ph.D. candidate of Operations Research, for his valuable suggestions and advice during the course of this investigation and computation in the preparation of this thesis.

I also wish to thank Ms. M. N. Marks for her very competent typing of this thesis.

I. INTRODUCTION

Properties of the stationary sequence of positive random variables $\{X_i\}$ which are formed from an independent and identically distributed exponential sequence $\{\varepsilon_i\}$ according to the linear model

$$X_i = \begin{cases} \beta\xi_i & \text{with probability } \beta, \\ & \qquad\qquad (0 \le \beta \le 1; i=0,\pm 1,\pm 2,\dots) \\ \beta\xi_i + \xi_{i+1} & \text{with probability } 1-\beta. \end{cases}$$

were discussed by Lawrance and Lewis [Ref. 1]. They gave a fairly complete picture of this model, and called it the EMA1 (Exponential Moving Average of order 1) process. It is clear that the adjacent elements of this sequence are correlated, but that the dependence is no greater than order one, i.e. X_i is independent of X_{i+2}, X_{i+3}, ... and so forth for pairs and triples.

In this paper, methods of estimating β and the properties of the estimates of β will be discussed, and then the properties of an analogous second order process are investigated. The new process, called the EMA2 model, is a sequence of positive random variables $\{X_i\}$ defined by

$$X_i = \begin{cases} \beta_2\xi_i & \text{w.p. } \beta_2; \\ \beta_2\xi_i + \beta_1\xi_{i+1} & \text{w.p. } (1-\beta_2)\beta_1 (0 \le \beta_2, \beta_1 \le 1; i=0,\pm 1,\dots); \\ \beta_2\xi_i + \beta_1\xi_{i+1} + \xi_{i+2} & \text{w.p. } (1-\beta_2)(1-\beta_1). \end{cases} \qquad (1.1)$$

The purpose in the creation of this model is to provide models for data with longer dependencies than that obtained with the first-order

model and to examine any tendencies of the upper bound on the serial correlations to increase. For the EMA1 model $0 \le \rho_1 \le 1/4$ and $\rho_k = 0$ for $k=2,3,\ldots\ldots$. For the EMA2 model it is shown that $0 \le \rho_1, \rho_2 \le 1/4$ and $\rho_k = 0$ for $k=3,4,\ldots$ In fact, the $\{X_i\}$ form a sequence of exponential random variables, and it will be seen from (1.1) that the successive elements X_i, X_{i+1}, X_{i+2} will be correlated. This model is also an alternative model to a renewal process.

The EMA2 model is shown to be a stationary point process. Distribution of the sums of X_i are discussed, and the joint distributions of two adjacent intervals X_i are derived and appear to be new bivariate exponential distributions. Extensions of the model and estimation problems are briefly discussed.

In developing the properties of the process, the similarities to a backward second order moving average which is defined as

$$X_i = \begin{cases} \beta_2 \xi_i & \text{w.p. } \beta_2; \\ \beta_2 \xi_i + \beta_1 \xi_{i-1} & \text{w.p. } (1-\beta_2)\beta_1 \quad (0 \leq \beta_2, \beta_1 \leq 1; i=0,\pm 1,\ldots.); \\ \beta_2 \xi_i + \beta_1 \xi_{i-1} + \xi_{i-2} & \text{w.p. } (1-\beta_2)(1-\beta_1). \end{cases} \tag{1.2}$$

will also be pointed out. Properties of the processes are very similar, but those of the forward model (1.1) have simpler derivations.

II. A BRIEF REVIEW OF THE EMA1 PROCESS

The EMA1 model is a stationary point process with exponential marginal distribution of the intervals $\{X_i\}$. Further X_i is dependent on X_{i-1} and X_{i+1}, but independent of all others, so the correlation $\rho_1 = \text{corr}(X_i, X_{i+1}) = \beta(1-\beta)$, $\rho_k=0$ for $k=2,3,\ldots.$

The Laplace transform of the p.d.f. of $T_r=X_1+X_2+\ldots+X_r$ is

$$\phi_r(s)=\frac{\lambda}{\lambda+s}\left\{\frac{\lambda(\lambda+2\beta s)}{(\lambda+\beta s)[\lambda+(1+\beta)s]}\right\}^{r-1} \qquad r\geq 1 \qquad (2.1)$$

Let $N_t^{(f)}$ be the number of events occurring in the interval $(0,t]$ beginning at an arbitrary event; and let $F_r(t)$ denote the distribution of T_r; then

$$\text{Prob}\{N_t^{(f)}=r\}=F_r(t)-F_{r+1}(t). \qquad r\geq 0$$

with $F_0(t)\equiv 1$ for $t\geq 0$. The p.d.f. of $N_t^{(f)}$ gives the generating function as

$$E[z^{N_t^{(f)}}]=\psi_f(z;t)=\sum_{r=0}^{\infty} z^r[F_r(t)-F_{r+1}(t)]=1+(z-1)\sum_{r=1}^{\infty} z^{r-1}F_r(t). \quad (2.2)$$

Inserting (2.1) in the Laplace transform of (2.2) gives

$$\psi_f^*(z;s)=\frac{\beta(1+\beta)s^2+[-\beta(1-\beta)z+2\beta+1]\lambda s+\lambda^2}{(s+\lambda)[\beta(1+\beta)s^2+(1+2\beta-2\beta z)\lambda s+(1-z)\lambda^2]} \qquad (2.3)$$

Differentiating (2.3) with respect to z, then setting $z=1$, gives the Laplace transform of the intensity function $m_f(t)$, as

$$m_f^*(z;s)=\frac{\lambda(\lambda+\beta s)[\lambda+(1+\beta)s]}{\beta(1+\beta)s(\lambda+s)[s+\lambda/(\beta^2+\beta)]}, \tag{2.4}$$

and inverting (2.4) gives

$$m_f(t)=\begin{cases}\lambda\{1+\frac{\beta(1-\beta)}{\beta^2+\beta-1}[e^{-\lambda t/(\beta^2+\beta)}-e^{-\lambda t}]\} & (\beta^2+\beta\neq 1)\\ \lambda(1+\beta^3\lambda te^{-\lambda t}) & (\beta^2+\beta=1)\end{cases}$$

The joint distribution of X_i and X_{i+1} is a bivariate exponential. Using a double Laplace transform we get

$$f^{**}_{X_i,X_{i+1}}(s_1,s_2)=\psi(\beta s_1)[\beta\psi(\beta s_2)+(1-\beta)\psi(s_1+\beta s_2)][\beta+(1-\beta)\psi(s_2)]$$

$$=\frac{\lambda^2(\lambda+\beta s_1+\beta s_2)}{(\lambda+\beta s_1)(\lambda+s_2)(\lambda+s_1+\beta s_2)}, \tag{2.5}$$

and using triple Laplace transform gives

$$f^{***}_{X_{i-1},X_i,X_{i+1}}(s_1,s_2,s_3)$$

$$=\psi(\beta s_1)[\beta\psi(\beta s_2)+(1-\beta)\psi(s_1+\beta s_2)][\beta\psi(\beta s_3)+(1-\beta)\psi(s_2+\beta s_3)][\beta+(1-\beta)\psi(s_3)]$$

Differentiating (2.5) with respect to s_2, setting $s_2=0_+$, and inverting with respect to s_1 and then dividing by the marginal (exponential) density of X_{i-1}, gives

$$E(X_i|X_{i-1}=t)=\lambda^{-1}[\beta\lambda t+\frac{1-2\beta}{1-\beta}+\frac{\beta}{1-\beta}e^{-\lambda(1-\beta)t/\beta}].$$

Similarly,

$$E(X_i | X_{i+1}=t)=\lambda^{-1}[1+\beta-e^{-(1-\beta)\lambda t/\beta}].$$

The two conditional variances are given by

$$\mathrm{Var}(X_i | X_{i-1}=t)=\lambda^{-2}\left[\frac{1-2\beta+2\beta^3}{(1-\beta)^2}+\frac{2\beta^2(1+\lambda t)}{1-\beta}e^{-(1-\beta)\lambda t/\beta}-\frac{\beta^2}{(1-\beta)^2}e^{-2(1-\beta)\lambda t/\beta}\right]$$

$$\mathrm{Var}(X_i | X_{i+1}=t)=\lambda^{-2}\left\{\frac{1+\beta+\beta^2-\beta^3}{1-\beta}-2\left[\frac{\beta}{1-\beta}+\frac{\lambda t}{\beta}\right]e^{-\lambda(1-\beta)t/\beta}-e^{-2(1-\beta)\lambda t/\beta}\right\}.$$

III. ESTIMATING β IN THE EMA1 MODEL

The EMA1 model is not time-reversible, and this comes out clearly in higher order joint moments. The results lead to a method for estimating β in the EMA1 model.

Define
$$C_{1,2}(k)=E(X_i X_{i+k}^2)-E(X_i)E(X_{i+k}^2),$$
$$C_{2,1}(k)=E(X_i^2 X_{i+k})-E(X_i^2)E(X_{i+k}),$$

which when $k=1$ gives
$$C_{1,2}(1)=E(X_i X_{i+1}^2)-E(X_i)E(X_{i+1}^2),$$
$$C_{2,1}(1)=E(X_i^2 X_{i+1})-E(X_i^2)E(X_{i+1}).$$

By the construction of EMA1, we have:

$$X_i^2=\begin{cases}\beta^2\varepsilon_i^2 & \text{w.p. } \beta^2, \\ \beta^2\varepsilon_i^2+2\beta\varepsilon_i\varepsilon_{i+1}+\varepsilon_{i+1}^2 & \text{w.p. } (1-\beta)^2.\end{cases} \qquad (0\le\beta\le 1;\ i=0,\pm1,\pm2,\ldots)$$

Hence, using straightforward combination, we get the joint expectation of X_i^2 and X_{i+1} as

$$\begin{aligned}E(X_i^2X_{i+1})=&E(\beta^3\varepsilon_i^2\varepsilon_{i+1})\beta^3+E(\beta^3\varepsilon_i^2\varepsilon_{i+1}+\beta^2\varepsilon_i^2\varepsilon_{i+2})\beta^2(1-\beta)\\ &+E(\beta^3\varepsilon_i^2\varepsilon_{i+1}+2\beta^2\varepsilon_i\varepsilon_{i+1}^2+\beta\varepsilon_{i+1}^3)\beta(1-\beta)^2\\ &+E(\beta^3\varepsilon_i^2\varepsilon_{i+1}+2\beta^2\varepsilon_i\varepsilon_{i+1}^2+\beta^2\varepsilon_i^2\varepsilon_{i+2}+2\beta\varepsilon_i\varepsilon_{i+1}\varepsilon_{i+2}+\beta\varepsilon_{i+1}^3+\varepsilon_{i+1}^2\varepsilon_{i+2})(1-\beta)^3.\end{aligned}$$

Simplification of this result leads to

$E(X_i^2X_{i+1})=\frac{1}{\lambda^3}(2+4\beta-2\beta^2-2\beta^3)$ which implies that $C_{2,1}(1)=\frac{2}{\lambda^3}\beta(1-\beta)(2+\beta)$.

Similarly, we get

$E(X_i X_{i+1}^2) = \frac{1}{\lambda^3}(2+2\beta-2\beta^3)$ which implies that $C_{1,2}(1) = \frac{2\beta}{\lambda^3}(1-\beta)(1+\beta)$.

Therefore, if we let

$$r = \frac{C_{2,1}(1)}{C_{1,2}(1)} = (2+\beta)/(1+\beta), \tag{3.1}$$

we have a function of β which decreases monotonically from 2 when $\beta \to 0$, to 3/2 when $\beta \to 1$. Thus there is a unique solution for β for any given r; note that when β is 0 or 1, the ratio is not defined. Solving (3.1) we get

$$\beta = (2-r)/(r-1) = \frac{2 - \frac{C_{2,1}(1)}{C_{1,2}(1)}}{\frac{C_{2,1}(1)}{C_{1,2}(1)} - 1} = \frac{2C_{1,2}(1) - C_{2,1}(1)}{C_{2,1}(1) - C_{1,2}(1)},$$

For estimating β, define

$$\hat{C}_{1,2}(1) = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i X_{i+1}^2 - (\bar{x})(\bar{x}^2),$$

$$\hat{C}_{2,1}(1) = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i^2 X_{i+1} - (\bar{x}^2)(\bar{x}),$$

Thus
$$\hat{\beta} = \frac{2\hat{C}_{1,2}(1) - \hat{C}_{2,1}(1)}{\hat{C}_{2,1}(1) - \hat{C}_{1,2}(1)}.$$

Now we check all the estimators, to see if they are asymptotically unbiased or not.

1. $$E[\hat{C}_{1,2}(1)] = \frac{1}{n-1}\sum_{i=1}^{n-1} E(X_i X_{i+1}^2) - E\left(\frac{1}{n^2}\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i^2\right).$$

Examining the estimate of the product of the means we have

$$\sum_{i=1}^{n} X_i \cdot \sum_{i=1}^{n} X_i^2 = (X_1+X_2+\ldots+X_n)(X_1^2+X_2^2+\ldots+X_n^2)$$
$$= nX_i^3+2(n-1)(X_iX_{i+1}^2+X_i^2X_{i+1})+(n-1)(n-2)X_iX_{i\pm 2}\,.$$

Thus

$$E\left(\frac{1}{n^2}\sum_{i=1}^{n}X_i\sum_{i=1}^{n}X_i^2\right)=\frac{1}{n^2}\frac{1}{\lambda^3}[6n+2(n-1)(4+6\beta-2\beta^2-4\beta^3)+2n^2-6n+4]$$
$$=\frac{2}{\lambda^3}+\frac{4}{n}(2+3\beta-\beta^2-2\beta^3)+\frac{4}{n^2}$$

and when $n\to\infty$, $E(\bar{X}\bar{X^2})\to 2/\lambda^3$.

Thus, since the estimators of $E(X_iX_{i+1}^2)$ and of $E(X_i^2X_{i+1})$ are unbiased, we get that asymptotically,

$$E[\hat{C}_{1,2}(1)]=\frac{2}{\lambda^3}\beta(1-\beta)(1+\beta);$$

$$E[\hat{C}_{2,1}(1)]=\frac{2}{\lambda^3}\beta(1-\beta)(2+\beta).$$

i.e. Both of these are unbiased estimators when n is large.

2. We now look at the ratio estimator of β , namely $\hat{\beta}$ to see if it is asymptotically unbiased. Note that the denominator in the expression for $\hat{\beta}$ is identically zero if $\beta=0$ or $\beta=1$, so that in what follows we assume that $0<\beta<1$. Let

$2\hat{C}_{1,2}(1)-\hat{C}_{2,1}(1)=Y$ and let

$\hat{C}_{2,1}(1)-\hat{C}_{1,2}(1)=Z$. Then

$E(Y)=2E[\hat{C}_{1,2}(1)]-E[\hat{C}_{2,1}(1)]=\frac{1}{\lambda^3}(2\beta^2-2\beta^3)=\mu_y$; and

$E(Z)= E[\hat{C}_{2,1}(1)]-E[\hat{C}_{1,2}(1)]=\frac{1}{\lambda^3}(2\beta-2\beta^2)=\mu_z$.

Now we can write $(Y/Z)=\frac{\mu_y}{\mu_z}\frac{Y/\mu_y}{Z/\mu_z}=\frac{\mu_y}{\mu_z}Y'(1+\frac{Z-\mu_z}{\mu_z})^{-1}$ where $Y'=Y/\mu_y$, so that

$$E(Y/Z)=\frac{\mu_y}{\mu_z}\{E(Y')-E[Y'(\frac{Z-\mu_z}{\mu_z})]+E[Y'(\frac{Z-\mu_z}{\mu_z})^2]-\ldots\}; \qquad (3.1.)$$

we assume that conditions for this expansion to hold as $n\to\infty$ are met.

Since $\mu_y/\mu_z=\beta$ and $E(Y')=1$, if $\hat{\beta}=Y/Z$ is to be unbiased, we must show the rest of the terms in (3.1) are all zeros.

Thus look at

$$E[Y'(Z-\mu_z)/\mu_z]=E(YZ)/\mu_y\mu_z-1.$$

we have

$$\begin{aligned}(YZ)&=[2\hat{C}_{1,2}(1)-\hat{C}_{2,1}(1)][\hat{C}_{2,1}(1)-\hat{C}_{1,2}(1)]\\
&=(n-1)^{-2}[2\sum_{i=1}^{n-1}X_iX_{i+1}^2-\sum_{i=1}^{n-1}X_i^2X_{i+1}][\sum_{i=1}^{n-1}X_i^2X_{i+1}-\sum_{i=1}^{n-1}X_iX_{i+1}^2]\\
&\quad-(\bar{x})(\overline{x^2})(\sum_{i=1}^{n-1}X_i^2X_{i+1}-\sum_{i=1}^{n-1}X_iX_{i+1}^2)/(n-1)\\
&=[(2U-W)(W-U)]/(n-1)^2-[\sum_{i=1}^{n}X_i\sum_{i=1}^{n}X_i^2(W-U)]/[n^2(n-1)],\end{aligned}$$

where $U=\sum_{i=1}^{n-1}X_iX_{i+1}^2=\sum_{i=1}^{n-1}U_i$, i.e. $U_i=X_iX_{i+1}^2$,

and $W=\sum_{i=1}^{n-1}X_i^2X_{i+1}=\sum_{i=1}^{n-1}W_i$, i.e. $W_i=X_i^2X_{i+1}$.

In addition $(2U-W)(W-U)=3UW-2U^2-W^2$.

Further we get

$$\begin{aligned}UW=&\sum_{i=1}^{n-1}U_iW_i+\sum_{i=3}^{n-1}U_iW_{i-2}+\sum_{i=2}^{n-1}U_iW_{i-1}+\sum_{i=1}^{n-2}U_iW_{i+1}\\
&+\sum_{i=1}^{n-3}U_iW_{i+2}+(n^2-7n+12)U_iW_{i\pm3},\end{aligned}$$

$$U^2=\sum_{i=1}^{n-1}U_i^2+2\sum_{i=1}^{n-2}U_iU_{i+1}+2\sum_{i=1}^{n-3}U_iU_{i+2}+(n-3)(n-4)U_iU_{i\pm3},$$

$$W^2=\sum_{i=1}^{n-1}W_i^2+2\sum_{i=1}^{n-2}W_iW_{i+1}+2\sum_{i=1}^{n-3}W_iW_{i+2}+(n-3)(n-4)W_iW_{i\pm3}.$$

It can be shown that all these joint expectations have finite expected value so, when $n\to\infty$, those terms only with coefficients n will go to zero when multiplying by $(\frac{1}{n-1})^2$. Thus asymptotically,

$$E[(2U-W)(W-U)/(n-1)^2]=E(3UW-2U^2-W^2)/(n-1)^2 \xrightarrow{n\to\infty}$$

$$3E(U_i)E(W_{i+3})-2E(U_i)E(U_{i+3})-E(W_i)E(W_{i+3})=(4\beta-4\beta^2+4\beta^3-8\beta^4+4\beta^5)/\lambda^6,$$

since $E(U_i)=E(X_iX_{i+1}^2)=(2+2\beta-2\beta^3)/\lambda^3$,

and $E(W_i)=E(X_i^2X_{i+1})=(2+4\beta-2\beta^2-2\beta^3)/\lambda^3$,

Similarly, $E[\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i^2 (W-U)/(n^3-n)] \xrightarrow{n\to\infty}$

$$E(X_i)E(X_i^2)[E(W_i)-E(U_i)]=(4\beta-4\beta^2)/\lambda^6,$$

Hence, $E(YZ)=\frac{1}{\lambda^6}(4\beta^3-8\beta^4+4\beta^5)$, when n is large.

Asymptotically,

$$E[Y'(Z-\mu_z)/\mu_z]=(4\beta^3-8\beta^4+4\beta^5)/(\mu_y\mu_z\lambda^6)-1=0.$$

In (3.1), the rest of the terms in the braces will also approach zero when n is large, so $E(\hat{\beta})=\beta$, i.e. $\hat{\beta}$ is an unbiased estimator.

An alternative way to estimate β is to use

$$\hat{\hat{\beta}}=[\lambda^3\hat{C}_{1,2}(1)]/2\hat{\rho}_1-1$$

where $\hat{\rho}_1$ is an estimator of ρ_1 the first order serial correlation of EMA1, and $\rho_1=\beta(1-\beta)$. [Ref. 1, p.5]

Define

$$\hat{\rho}_1=\lambda^2[\sum_{i=1}^{n-1} X_iX_{i+1}/(n-1)-(\bar{x})^2]$$

Then

$$E(\hat{\rho}_1)=\lambda^2\{\sum_{i=1}^{n-1} E(X_iX_{i+1})/(n-1)-E[(\sum_{i=1}^{n} X_i)^2]/n^2\}.$$

Again using the same argument as above, we have

$$E(X_iX_{i+1})=(1+\beta-\beta^2)/\lambda^2.$$

Also $$\left(\sum_{i=1}^{n} X_i\right)^2=\sum_{i=1}^{n} X_i^2+2\sum_{i=1}^{n-1} X_iX_{i+1}+2\sum_{i=1}^{n-2}\sum_{j=2}^{n-i} X_iX_{i+j},$$

so that $$E[(\bar{x})^2]=n^{-2}[2n/\lambda^2+2(n-1)(1+\beta-\beta^2)/\lambda^2$$

$$+(n-1)(n-2)/\lambda^2]\xrightarrow{n\to\infty} 1/\lambda^2.$$

Consequently $E(\hat{\rho}_1)=\lambda^2\{(n-1)(1+\beta-\beta^2)/[\lambda^2(n-1)]-E[(\bar{x})^2]\}\xrightarrow{n\to\infty}\beta(1-\beta)$.

Thus $\hat{\rho}_1$ is an unbiased estimator for ρ_1 when n is large.

Now assume $n\to\infty$ and let $Y=\lambda^3\hat{C}_{1,2}(1)$ and $Z=2\hat{\rho}_1$. Thus

$$E(Y)=2\beta(1-\beta)(1+\beta),\quad \text{and} \quad E(Z)=2\beta(1-\beta),$$

$$\mu_y/\mu_z=1+\beta,\quad \mu_y\mu_z=4(\beta^2-\beta^3-\beta^4+\beta^5),$$

and by the expansion used above

$$E(Y/Z)=\frac{\mu_Y}{\mu_Z}\{E(y')-E[Y'(\frac{Z-\mu_Z}{\mu_Z})]+E[Y'(\frac{Z-\mu_Z}{\mu_Z})^2]-\ldots.\} \qquad (3.2)$$

We want to show that the terms in (3.2) beyond the first are zero, so we look at

$$E[Y'(Z-\mu_z)/\mu_z]=E(YZ)/\mu_y\mu_z-1.$$

We have

$$(YZ)=2\lambda^3\hat{\rho}_1\hat{C}_{1,2}(1)=2\lambda^5[\sum_{i=1}^{n-1} X_iX_{i+1}/(n-1)-(\sum_{i=1}^{n} X_i/n)^2]$$

$$[\sum_{i=1}^{n-1} X_iX_{i+1}^2/(n-1)-\sum_{i=1}^{n} X_i\sum_{i=1}^{n} X_i^2/n^2].$$

Therefore

$$E(YZ) \xrightarrow{n\to\infty} 2\lambda^5[E(X_iX_{i+1})E(X_iX_{i+1}^2)-E(X_iX_{i+1})E(X_i)E(X_i^2)$$
$$-E(X_i)E(X_i)E(X_iX_{i+1}^2)+E(X_i)E(X_i)E(X_i)E(X_i^2)]$$
$$=4(\beta^2-\beta^3-\beta^4+\beta^5),$$

so that,

$$E[Y'(Z-\mu_z)/\mu_z]=E(YZ)/\mu_y\mu_z-1=1-1=0.$$

Similarly, we can show the rest of the terms in the braces of (3.2) all approach zero when n is large. Hence

$$E(\hat{\hat{\beta}})=E(Y/Z)-1=1+\beta-1=\beta,$$

i.e. $\hat{\hat{\beta}}$ is also an unbiased estimator of β when n is large.

IV. COMPARISON OF THE ESTIMATORS

It has been shown in the last section that the two estimators $\hat{\beta}$, $\hat{\hat{\beta}}$ are unbiased asymptotically, provided that $\beta \neq 0$, or $\beta \neq 1$, but when the sample size n is not large enough, the bias term should be considered. For simplification, any finite term divided by the second or higher power of n will be neglected.

For $\hat{\beta}$, let $Y=2\hat{C}_{1,2}(1)-\hat{C}_{2,1}(1)$; $Z=\hat{C}_{2,1}(1)-\hat{C}_{1,2}(1)$. All the estimators here are defined as before. Thus

$$E(Y)=2\beta^2(1-\beta)/\lambda^3-4(2+3\beta-\beta^2-2\beta^3)/n\lambda^3=\mu_y;$$

$$E(Z)=2\beta(1-\beta)/\lambda^3=\mu_z;$$

$$E(Y/Z)=\mu_y\mu_z^{-1}\{1-[E(YZ)/\mu_y\mu_z-1]+\ldots.\} .$$

Using those results listed in APPENDIX B and neglecting the higher power terms, we have

$$E(YZ)=E(Y)E(Z)+(-12-560\beta+1472\beta^2-1148\beta^3+532\beta^4-1120\beta^5$$
$$+740\beta^6+248\beta^7-164\beta^8)/\lambda^6 n .$$

Hence $E(\hat{\beta})=[2\mu_y\mu_z-E(YZ)]/\mu_z^2$

$$=\beta+\frac{1}{4\beta^2(1-\beta)^2}\cdot\frac{1}{n}(12+528\beta-1488\beta^2+1212\beta^3-516\beta^4$$
$$+1088\beta^5-740\beta^6-248\beta^7+164\beta^8). \quad (4.1)$$

For $\hat{\hat{\beta}}$, let $Y=\lambda^3\hat{C}_{1,2}(1)$; $Z=2\hat{\rho}_1$.

Again all the estimators here are defined as before. Thus

$$E(Y)=2\beta(1-\beta)(1+\beta)-4(2+3\beta-\beta^2-2\beta^3)/n=\mu_y, \quad \text{and}$$

$$E(Z)=2\beta(1-\beta)+(2+4\beta-4\beta^2)/n=\mu_z. \tag{4.2}$$

In (4.2), the maximum value of $(2+4\beta-4\beta^2)$ occurs at $\beta=1/2$, and equals to 3, when divided by n, it can be neglected, so $\mu_z=2\beta(1-\beta)$. Similarly as above we have

$$E(YZ)=E(Y)E(Z)+(4-16\beta+34\beta^2+24\beta^3-26\beta^4+8\beta^5-24\beta^6)/n ,$$

and consequently

$$E(\hat{\hat{\beta}})=\beta+\frac{1}{4\beta^2(1-\beta)^2}\,\frac{1}{n}(-4+16\beta-22\beta^2-32\beta^3+14\beta^4+24\beta^6). \tag{4.3}$$

Now compare (4.1) and (4.3); both of them are divided by $4\beta^2(1-\beta)^2$. When β approaches 0 or 1, the values of bias term will be very large, though both of the sums of the coefficients of the β's in the parentheses are zero when β approaches 0 or 1. Figures 1 to 4 give the shape of the curves of $f(\beta)$ and bias for different values of β. From the figures it is obvious that $\hat{\hat{\beta}}$ is better than $\hat{\beta}$.

The variances of those estimators are very messy for hand computation, and have not been worked out for this thesis.

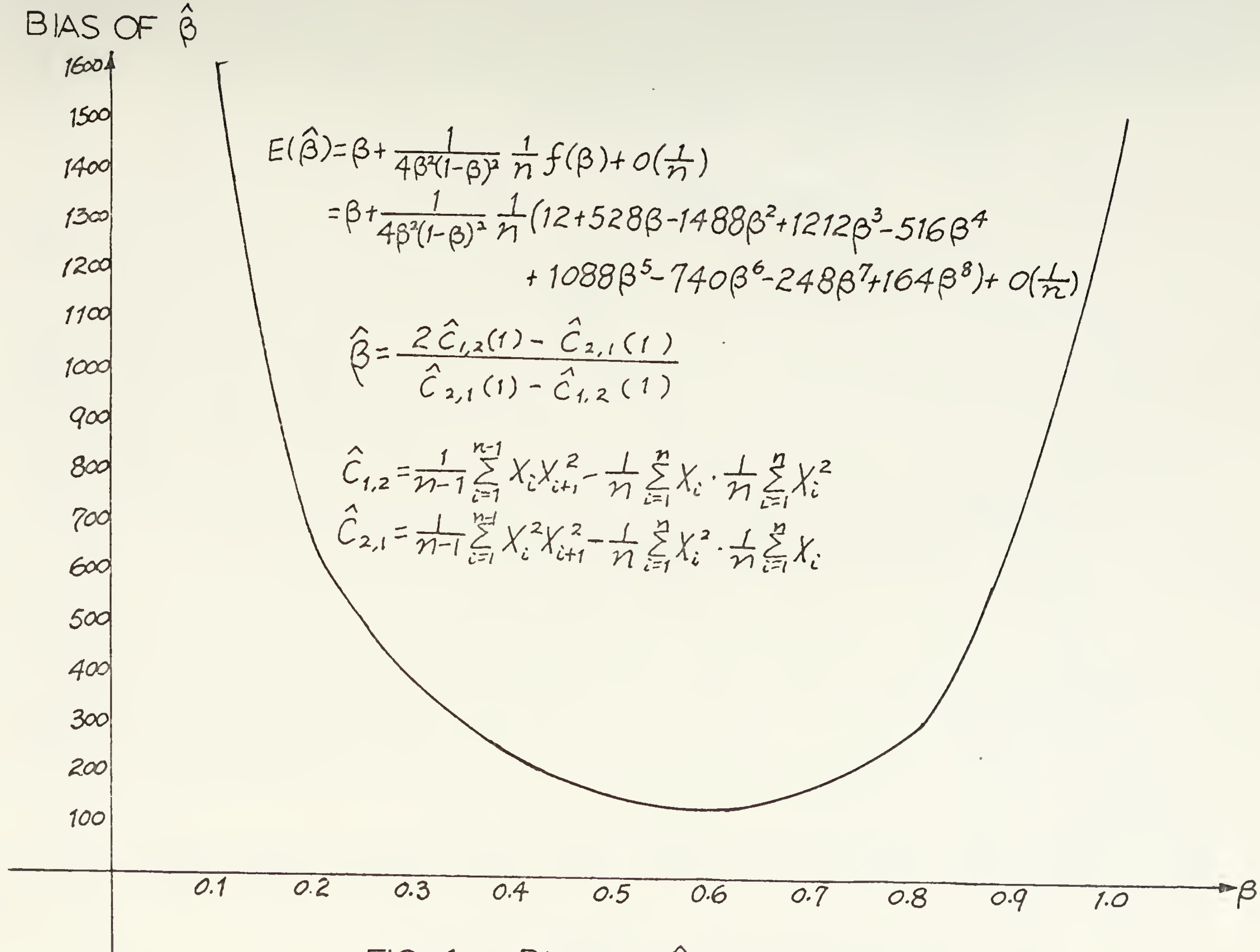


FIG. 1 BIAS OF $\hat{\beta}$ VS β

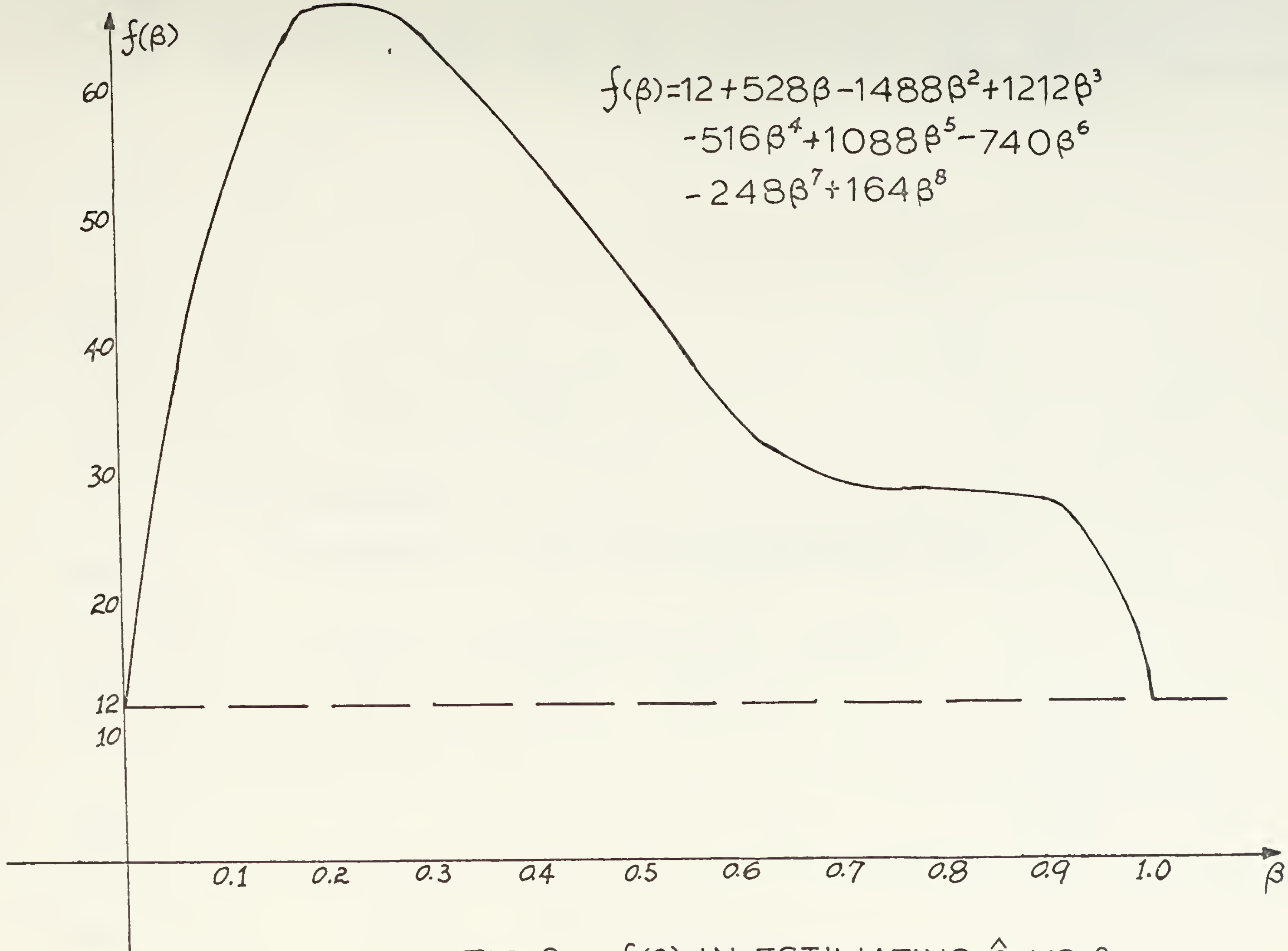
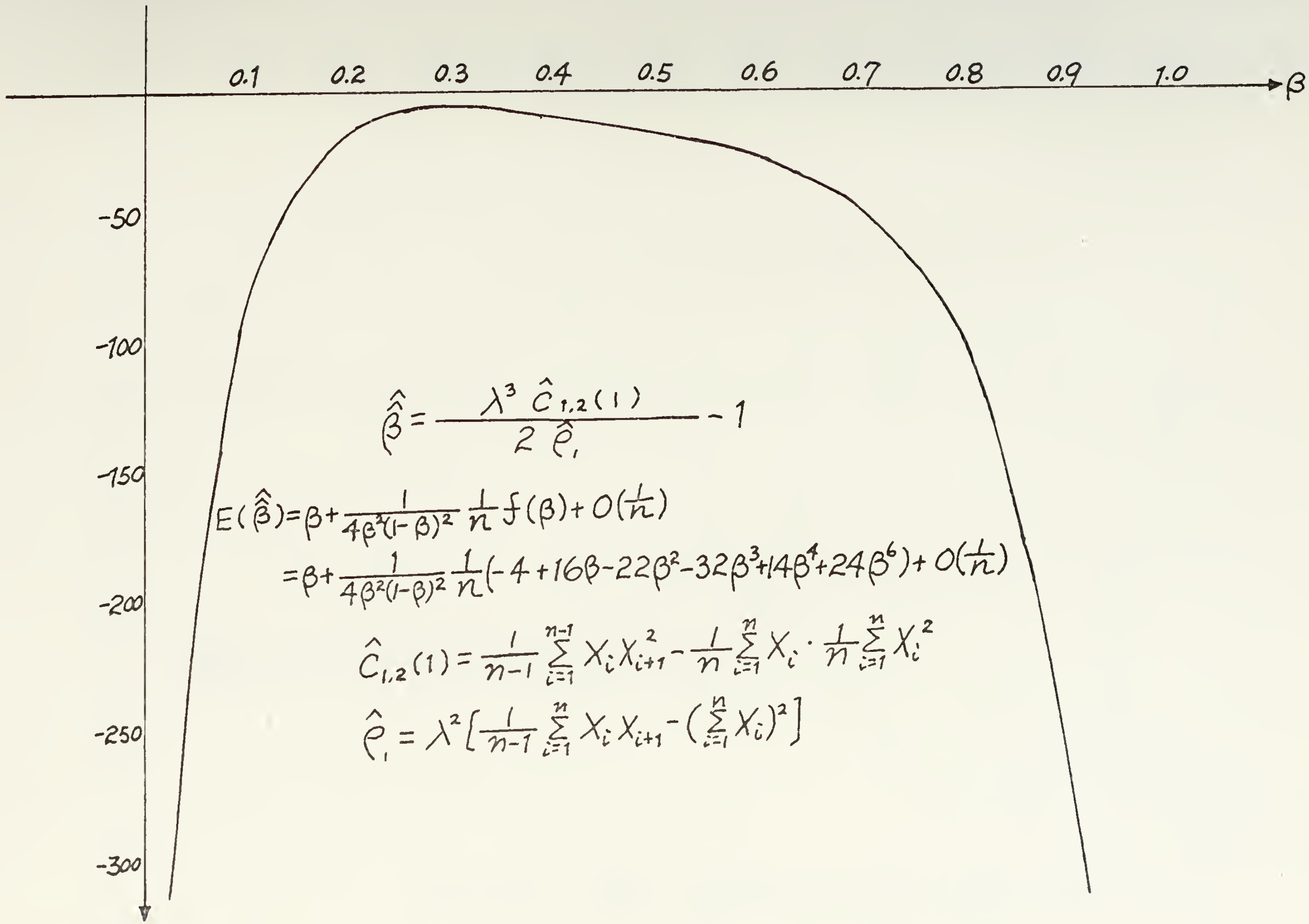


FIG. 2 $f(\beta)$ IN ESTIMATING $\hat{\beta}$ VS β



$$\hat{\hat{\beta}} = \frac{\lambda^3 \hat{C}_{1,2}(1)}{2\hat{\rho}_1} - 1$$

$$E(\hat{\hat{\beta}}) = \beta + \frac{1}{4\beta^2(1-\beta)^2}\frac{1}{n} f(\beta) + O(\tfrac{1}{n})$$

$$= \beta + \frac{1}{4\beta^2(1-\beta)^2}\frac{1}{n}(-4 + 16\beta - 22\beta^2 - 32\beta^3 + 14\beta^4 + 24\beta^6) + O(\tfrac{1}{n})$$

$$\hat{C}_{1,2}(1) = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i X_{i+1}^2 - \frac{1}{n}\sum_{i=1}^{n} X_i \cdot \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\hat{\rho}_1 = \lambda^2\left[\frac{1}{n-1}\sum_{i=1}^{n} X_i X_{i+1} - \left(\sum_{i=1}^{n} X_i\right)^2\right]$$

BIAS OF $\hat{\hat{\beta}}$

FIG. 3 BIAS OF $\hat{\hat{\beta}}$ VS β

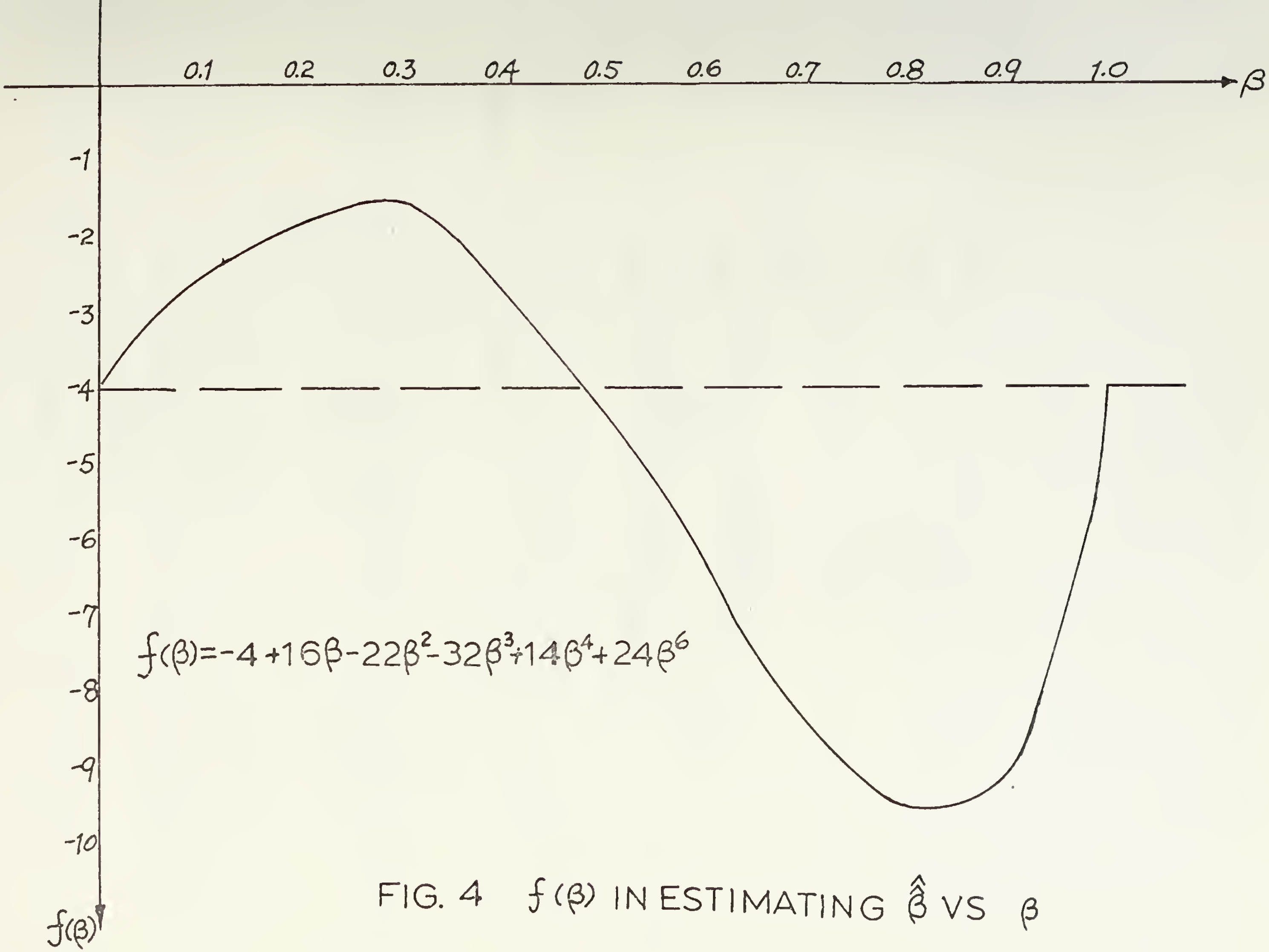


FIG. 4 $f(\beta)$ IN ESTIMATING $\hat{\hat{\beta}}$ VS β

V. SOME BASIC ASPECTS OF THE EMA2 MODEL

The simplest aspect of the EMA2 model is the exponential marginal distribution of the intervals $\{X_i\}$; in point process terminology [Ref. 2] this is the synchronous distribution of intervals and refers to the distribution of the interval from an arbitrarily chosen event to the next two events. For the Laplace transform of its probability density function (p.d.f.) $f_{X_i}(x)$, we write

$$
\begin{aligned}
f^*_{X_i}(s) &= E\{e^{-sX_i}\} \\
&= E\{e^{-s\beta_2\varepsilon_i}\}\beta_2 + E\{e^{-s\beta_2 - s\beta_1\varepsilon_{i+1}}\}(1-\beta_2)\beta_1 \\
&\quad + E\{e^{-s\beta_2\varepsilon_i - s\beta_1\varepsilon_{i+1} - s\varepsilon_{i+2}}\}(1-\beta_2)(1-\beta_1)
\end{aligned}
\tag{5.1}
$$

using (1.1). Since the i.i.d. random variable ε_i have exponential distributions with parameters λ , their Laplace transform is $\lambda/(\lambda+s)$. Thus (5.1) becomes

$$
\begin{aligned}
f^*_{X_i}(s) &= \frac{\lambda}{\lambda+\beta_2 s}\cdot\beta_2 + \frac{\lambda}{\lambda+\beta_2 s}\,\frac{\lambda}{\lambda+\beta_1 s}(1-\beta_2)\beta_1 + \frac{\lambda}{\lambda+\beta_2 s}\,\frac{\lambda}{\lambda+\beta_1 s}\,\frac{\lambda}{\lambda+s}(1-\beta_2)(1-\beta_1) \\
&= \frac{\lambda}{\lambda+s}.
\end{aligned}
$$

This demonstrates that the X_i have identical exponential distributions as asserted. The parameter λ is the number of events per unit time or the rate of the point process.

The correlation between X_i and X_{i+1} can be obtained on considering the product of X_i from (1.1) with

$$
X_{i+1} = \begin{cases}
\beta_2\varepsilon_{i+1} & \text{w.p. } \beta_2, \\
\beta_2\varepsilon_{i+1} + \beta_1\varepsilon_{i+2} & \text{w.p. } (1-\beta_2)\beta_1, \quad (0 \le \beta_2, \beta_1 \le 1;\ i = 0, \pm 1, \pm 2, \ldots) \\
\beta_2\varepsilon_{i+1} + \beta_1\varepsilon_{i+2} + \varepsilon_{i+3} & \text{w.p. } (1-\beta_2)(1-\beta_1).
\end{cases}
$$

Thus, again using straightforward conditioning arguments,

$$\begin{aligned}
E(X_i,X_{i+1}) = {} & E(\beta_2^2\varepsilon_i\varepsilon_{i+1})\beta_2^2 + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_i\varepsilon_{i+2})\beta_2\beta_1(1-\beta_2) \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_{i+1}^2)\beta_2\beta_1(1-\beta_2) \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_i\varepsilon_{i+2}+\beta_2\varepsilon_i\varepsilon_{i+3})\beta_2(1-\beta_1)(1-\beta_2) \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_{i+1}^2+\beta_2\varepsilon_{i+1}\varepsilon_{i+2})\beta_2(1-\beta_2)(1-\beta_1) \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_i\varepsilon_{i+2}+\beta_2\beta_1\varepsilon_{i+1}^2+\beta_1^2\varepsilon_{i+1}\varepsilon_{i+2})(1-\beta_2)^2\beta_1^2 \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_i\varepsilon_{i+2}+\beta_2\varepsilon_i\varepsilon_{i+3}+\beta_2\beta_1\varepsilon_{i+1}^2+\beta_1^2\varepsilon_{i+1}\varepsilon_{i+2} \\
& \quad +\beta_1\varepsilon_{i+1}\varepsilon_{i+3})(1-\beta_2)^2\beta_1(1-\beta_1) \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_i\varepsilon_{i+2}+\beta_1\beta_2\varepsilon_{i+1}^2+\beta_1^2\varepsilon_{i+1}\varepsilon_{i+2}+\beta_2\varepsilon_{i+1}\varepsilon_{i+2} \\
& \quad +\beta_1\varepsilon_{i+2}^2)(1-\beta_2)^2\beta_1(1-\beta_1) \\
& + E(\beta_2^2\varepsilon_i\varepsilon_{i+1}+\beta_2\beta_1\varepsilon_i\varepsilon_{i+2}+\beta_2\varepsilon_i\varepsilon_{i+3}+\beta_2\beta_1\varepsilon_{i+1}^2+\beta_1^2\varepsilon_{i+1}\varepsilon_{i+2} \\
& \quad +\beta_1\varepsilon_{i+1}\varepsilon_{i+3}+\beta_2\varepsilon_{i+1}\varepsilon_{i+2}+\beta_1\varepsilon_{i+2}^2+\varepsilon_{i+2}\varepsilon_{i+3})(1-\beta_2)^2(1-\beta_1)^2.
\end{aligned}$$

and simplification of this result leads to

$$\rho_1=\mathrm{corr}(X_i,X_{i+1})=\frac{\mathrm{cov}(X_i,X_{i+1})}{(\mathrm{var}X_i\cdot\mathrm{var}X_{i+1})^{\frac{1}{2}}}=\beta_1(1-\beta_2)\left[1-\beta_1(1-\beta_2)\right] \qquad (5.2)$$

Similarly we have

$$\rho_2=\mathrm{corr}(X_i,X_{i+2})=\beta_2(1-\beta_2)(1-\beta_1) \qquad (5.3)$$

By the construction of EMA2, $\rho_j=0$ for $j\geq 3$. The result for ρ_1 (5.2) equals zero when $\beta_2=0$ and $\beta_1=1$, or $\beta_2=1$, or $\beta_1=0$; and will approach its maximum value at $\beta_1(1-\beta_2)=1/2$. This occurs when $\beta_1=1/[2(1-\beta_2)]$. The result for ρ_2 (5.3) equals to zero when $\beta_2=0$, or $\beta_1=1$, or $\beta_2=1$; and will approach its maximum value at $\beta_2=1/2$ and $\beta_1=0$. Therefore, the serial correlations of EMA2 are all non-negative and bounded above by 1/4.

Now the stationarity of the EMA2 process will be discussed.

Define $$E(X_i)=m(i) \tag{5.4}$$
$$E[X_i-m(i)][X_{i+k}-m(i+k)]=\mathrm{Cov}(X_i,X_{i+k})=\sigma(i,i+k). \tag{5.5}$$

A stochastic process with a discrete time parameter is said to be "stationary" (or stationary in the strict sense) if the distribution of X_i, X_{i+1}, ..., X_{i+j} is the same as the distribution of X_{i+k}, X_{i+1+k}, ..., X_{i+j+k}, for every finite set of integers $\{1,2.....j\}$ and for every integer k. This definition is equivalent to requiring that the probability measure for the sequence $\{X_i\}$ be the same as that of $\{X_{i+k}\}$ for every integer k. If the first-order moments exist, stationarity implies that $E(X_i)=E(X_{i+k})$ for all $i,k=0,\pm 1,\pm 2,\ldots$ (5.6)
Since (X_i,X_{i+1}) has the same distribution as (X_{i+k},X_{i+1+k}), existence of the second-order moments and stationarity imply

$$\sigma(i,i+j)=\sigma(i+k,i+j+k)\ . \tag{5.7}$$

Setting $k=-i-1$ gives $\sigma(i,i+j)=\sigma[i-(i+j)]=\sigma(j)$. (5.8)

In the normal case properties (5.6) and (5.7) determine that the stochastic process is stationary.

A stochastic process is said to be stationary in the wide sense or weakly stationary or stationary of second order if the mean function and the covariance function defined as in (5.4) and (5.5) exist and satisfy the relations (5.6) and (5.7); i.e. the mean is a constant, independent of time, and the covariance of any two variables depends only on their distance apart in time. Obviously, any process which is stationary in the strict sense and has finite variance is also stationary in the wide sense. In the normal case discussed above stationary in the strict sense and in the wide sense are equivalent.

We have proved that the X_i have identical exponential distributions, which implies that $E(X_i)$ exists and $E(X_i)=E(X_{i+k})$ for all $i,k=0,\pm 1,\pm 2,\ldots$ Also we have

$$\mathrm{cov}(X_i,X_{i\pm 0})=1/\lambda^2$$

$$\mathrm{cov}(X_i,X_{i\pm 1})=\frac{1}{\lambda^2}\{\beta_1(1-\beta_2)-[\beta_1(1-\beta_2)]^2\},$$

$$\mathrm{cov}(X_i,X_{i\pm 2})=\frac{1}{\lambda^2}[\beta_2(1-\beta_2)(1-\beta_1)],$$

$$\mathrm{cov}(X_i,X_{i+k})=0. \qquad \text{for } k=3,4,5,\ldots\ldots$$

All these expectations and covariances are independent of time i, so we conclude that the EMA2 process is stationary in the wide sense.

The independent exponential sequences and EMA1 models are the special cases of the EMA2 model; these aspects of the EMA2 model are described in the following table:

Values of β_2 & β_1 in EMA2 model	When we set	X_i sequence reduces to
$\beta_2 = 0$; $\beta_1 = \beta_1$	$\beta_1 = \beta$; $\varepsilon_{i+1} = \varepsilon_i$	EMA1 model
$\beta_2 = \beta_2$; $\beta_1 = 1$	$\beta_2 = \beta$	EMA1 model
$\beta_2 = \beta_2$; $\beta_1 = 0$	$\beta_2 = \beta$; $\varepsilon_{i+2} = \varepsilon_{i+1}$ Now the adjacent elements are indep. if keep ε_{i+2} no change	EMA1 model
$\beta_2 = 1$; $\beta_1 = \beta_1$	$X_i = \varepsilon_i$ w.p. 1	Poisson process (i.i.d)
$\beta_2 = 0$; $\beta_1 = 1$	$X_i = \varepsilon_{i+1}$ w.p. 1	Poisson process (i.i.d)
$\beta_2 = 0$; $\beta_1 = 0$	$X_i = \varepsilon_{i+2}$ w.p. 1	Poisson process (i.i.d)

This gives checks on most of the results, for the serial correlations. In the 3rd case X_i and X_{i+1} are independent, so $\rho_1 = 0$; but X_i and X_{i+2} are dependent, so $\rho_2 = \beta_2(1-\beta_2)$, which is the same expression of ρ_1 in EMA1. The serial correlations in the last three cases are all zero, since all of them have i.i.d. elements.

Also, even the backward model (1.2) could be equally treated to produce similar but different results. However, there is no time-reversibility in the process, in the sense that $\{X_1, X_2, \ldots\ldots X_k\}$ does not have the same joint probability distribution as $\{X_{-1}, X_{-2}, \ldots\ldots X_{-k}\}$ for all finite k, where $k \geq 2$.

VI. DISTRIBUTION OF SUMS IN $\{X_i\}$ SEQUENCE OF THE EMA2 MODEL

In the point process theory of the model, the distribution of the sums $T_r=X_1+X_2+\ldots+X_r$ are very useful; if these can be obtained then the distribution of counts, both in the synchronous and asynchronous mode, can then be derived. It would, therefore, be a particularly attractive feature of the EMA2 model if the distribution of the T_r could be obtained. Unfortunately, it is not possible to get a simple expression of the Laplace transform of the p.d.f. of T_r as in EMA1.

However, a general derivation will now be given. Define $\psi(s)$ as the Laplace transform of the p.d.f. of the ε_i distribution; except where otherwise remarked this distribution is exponential of parameter λ and so $\psi(s)=\lambda/(\lambda+s)$. Define the triple Laplace transform of the p.d.f. of T_r, ε_{r+1} and ε_{r+2} as

$$\phi_r(s_1,s_2,s_3)=E\{e^{-s_1T_r-s_2\varepsilon_{r+1}-s_3\varepsilon_{r+2}}\} \qquad \text{for } r=1,2,\ldots\ldots \tag{6.1}$$

For $r=1$, we have

$$\begin{aligned}\phi_1(s_1,s_2,s_3)&=E\{e^{-s_1X_1-s_2\varepsilon_2-s_3\varepsilon_3}\}\\ &=E\{e^{-s_1\beta_2\varepsilon_1-s_2\varepsilon_2-s_3\varepsilon_3}\}\beta_2+E\{e^{-s_1(\beta_2\varepsilon_1+\beta_1\varepsilon_2)-s_2\varepsilon_2-s_3\varepsilon_3}\}(1-\beta_2)\beta_1\\ &\quad+E\{e^{-s_1(\beta_2\varepsilon_1+\beta_1\varepsilon_2+\varepsilon_3)-s_2\varepsilon_2-s_3\varepsilon_3}\}(1-\beta_1)(1-\beta_2)\\ &=\psi(\beta_2s_1)[\beta_2\psi(s_2)\psi(s_3)+(1-\beta_2)\beta_1\psi(\beta_1s_1+s_2)\psi(s_3)\\ &\quad+(1-\beta_2)(1-\beta_1)\psi(\beta_1s_1+s_2)\psi(s_1+s_3)]\end{aligned}$$

Now we relate $\phi_r(s_1,s_2,s_3)$ and $\phi_{r-1}(s_1,s_2,s_3)$ using the expression

$$T_r=T_{r-1}+X_r=\begin{cases} T_{r-1}+\beta_2\mathcal{E}_r & \text{w.p. } \beta_2 \\ T_{r-1}+\beta_2\mathcal{E}_r+\beta_1\mathcal{E}_{r+1} & \text{w.p. } (1-\beta_2)\beta_1, \\ T_{r-1}+\beta_2\mathcal{E}_r+\beta_1\mathcal{E}_{r+1}+\mathcal{E}_{r+2} & \text{w.p. } (1-\beta_2)(1-\beta_1). \end{cases}$$

Then we have

$$\begin{aligned}\phi_r(s_1,s_2,s_3)&=E\left\{e^{-s_1(T_{r-1}+\beta_2\mathcal{E}_r)-s_2\mathcal{E}_{r+1}-s_3\mathcal{E}_{r+2}}\right\}\beta_2 \\ &\quad+E\left\{e^{-s_1(T_{r-1}+\beta_2\mathcal{E}_r+\beta_1\mathcal{E}_{r+1})-s_2\mathcal{E}_{r+1}-s_3\mathcal{E}_{r+2}}\right\}(1-\beta_2)\beta_1 \\ &\quad+E\left\{e^{-s_1(T_{r-1}+\beta_2\mathcal{E}_r+\beta_1\mathcal{E}_{r+1}+\mathcal{E}_{r+2})-s_2\mathcal{E}_{r+1}-s_3\mathcal{E}_{r+2}}\right\}(1-\beta_2)(1-\beta_1) \\ &=\phi_{r-1}(s_1,\beta_2s_1,s_2)\psi(s_3)\beta_2 \\ &\quad+\phi_{r-1}(s_1,\beta_2s_1,\beta_1s_1+s_2)[\beta_1(1-\beta_2)\psi(s_3)+(1-\beta_2)(1-\beta_1)\psi(s_1+s_3)].\end{aligned}$$

Continuing, we can write

$$\begin{aligned}\phi_{r-1}(s_1,\beta_2s_1,s_2)&=\phi_{r-2}(s_1,\beta_2s_1,\beta_2s_1)\psi(s_2)\beta_2 \\ &\quad+\phi_{r-2}(s_1,\beta_2s_1,\beta_1s_1+\beta_2s_1)[\beta_1(1-\beta_2)\psi(s_2)+(1-\beta_2)(1-\beta_1)\psi(s_1+s_2)].\end{aligned}$$

$$\begin{aligned}\phi_{r-1}(s_1,\beta_2s_1,\beta_1s_1+s_2)&=\phi_{r-2}(s_1,\beta_2s_1,\beta_2s_1)\psi(\beta_1s_1+s_2)\beta_2 \\ &\quad+\phi_{r-2}(s_1,\beta_2s_1,\beta_1s_1+\beta_2s_1)\cdot \\ &\quad[\beta_1(1-\beta_2)\psi(\beta_1s_1+s_2)+(1-\beta_1)(1-\beta_2)\psi(s_1+\beta_1s_1+s_2)].\end{aligned}$$

and solve it recursively; the procedure is difficult by hand but could possibly be manipulated on a computer. Setting $s_2=0$ and $s_3=0$, we have the Laplace transform of the p.d.f. of T_r.

The first few sums have transforms as follows:

$\phi_1(s,0,0) = \frac{\lambda}{\lambda+s}$ which implies that $T_1 = X_1 \sim$ exponential (λ) as is expected. For T_2 we have:

$$\phi_2(s,0,0) = \frac{\lambda}{\lambda+s} \; \frac{\lambda[(\lambda+\beta_2 s)(\lambda+\beta_2 s+2\beta_1 s)+\beta_2\beta_1^2 s^2(2-\beta_2)]}{(\lambda+\beta_2 s)(\lambda+s+\beta_1 s)(\lambda+\beta_2 s+\beta_1 s)} \tag{6.2}$$

If we let $\beta_1 = 1$, (6.2) reduces to $\frac{\lambda}{\lambda+s}\,\frac{\lambda(\lambda+2\beta_2 s)}{(\lambda+\beta_2 s+s)(\lambda+\beta_2 s)}$ which is the Laplace transform of the p.d.f. of T_2 in EMA1. [Ref. 1, p.8]

If we let $\beta_1 = 0$, (6.2) reduces to $\left(\frac{\lambda}{\lambda+s}\right)^2$ which implies that T_2 is the sum of two independent exponential random variables.

For T_3 we have

$$\phi_3(s,0,0) = \frac{\lambda}{\lambda+s} \; \frac{\lambda^2[(A)(B)+(C)(D)]}{(\lambda+s+\beta_1 s+\beta_2 s)(\lambda+\beta_1 s+\beta_2 s)^2(\lambda+s+\beta_2 s)(\lambda+s+\beta_1 s)(\lambda+\beta_2 s)^2},$$

where $A = [s^2(2\beta_2^2+3\beta_1\beta_2+\beta_1^2\beta_2)+\lambda s(3\beta_2+2\beta_1)+\lambda^2](1-\beta_2)$ (6.3)

$$B = s^3(\beta_1\beta_2+\beta_1\beta_2^2+\beta_1^2\beta_2^2+\beta_1^2\beta_2)+\lambda s^2(\beta_2+\beta_2^2+3\beta_1\beta_2+\beta_1^2\beta_2+2\beta_1)+\lambda^2 s(1+2\beta_1+2\beta_2)+\lambda^3$$

$$C = s^2(\beta_2+\beta_2^2+\beta_1\beta_2^2+2\beta_1\beta_2+\beta_1^2\beta_2)+\lambda s(\beta_2^2+2\beta_2+2\beta_1\beta_2)+\beta_2\lambda^2$$

$$D = s^3(2\beta_2^3+4\beta_1\beta_2^2+2\beta_1\beta_2^2)+\lambda s^2(5\beta_2^2+6\beta_1\beta_2+\beta_1^2)+s^2\lambda(4\beta_2+2\beta_1)+\lambda^3.$$

If we let β_1, β_2 equal to zero or one, we have some interesting results.

When $\beta_2 = 0$, $(A)(B) = \lambda^2(\lambda+s)(\lambda+2\beta_1 s)^2$ and $(C)(D) = 0$, so that (6.3) reduces to

$$\frac{\lambda}{\lambda+s}\left[\frac{\lambda(\lambda+2\beta_1 s)}{(\lambda+\beta_1 s+s)(\lambda+\beta_1 s)}\right]^2$$

which is the Laplace transform of the p.d.f. of T_3 in EMA1 for $\beta = \beta_1$.

When

$$\beta_1=1,\ A=(1-\beta_2)(\lambda+\beta_2 s+2s)(\lambda+2\beta_2 s),\quad B=(\lambda+\beta_2 s)(\lambda+2s)(\lambda+s+\beta_2 s),$$

$$C=\beta_2(\lambda+2s)(\lambda+2s+\beta_2 s) \text{ and } D=(\lambda+2\beta_2 s)(\lambda+s+\beta_2 s)^2,$$

and this will give the same result as above for $\beta=\beta_2$.

When

$$\beta_2=1,\quad (A)(B)=0,\ (C)(D)=(\lambda+2s+\beta_1 s)(\lambda+s+\beta_1 s)^3(\lambda+2s),$$

(6.3) reduces to $\lambda^3/(\lambda+s)^3$, indicating that the $\{X_i\}$ sequence are i.i.d. exponentials.

When

$$\beta_1=0,\ A=(1-\beta_2)(\lambda+\beta_2 s)(\lambda+2\beta_2 s),\ B=\lambda(\lambda+\beta_2 s)(\lambda+s+\beta_2 s),$$

$$C=\beta_2(\lambda+s)(\lambda+s+\beta_2 s) \text{ and } D=(\lambda+\beta_2 s)^2(\lambda+2\beta_2 s)$$

so that (6.3) reduces to

$$\left(\frac{\lambda}{\lambda+s}\right)^2\frac{\lambda(\lambda+2\beta_2 s)}{(\lambda+\beta_2 s)(\lambda+s+\beta_2 s)},$$

which means X_1 and X_3 form an EMA1 model, X_2 is exponential (λ) and independent of X_1 and X_3 .

VII. THE JOINT DISTRIBUTION OF X_i AND X_{i+1} IN EMA2

We now discuss the joint distribution of X_i and X_{i+1} which will be a bivariate exponential distribution. Several authors have discussed bivariate exponential distributions, including Downton (1970), who makes some comparisons with those of Gumbel, Moran and Marshall-Olkin. The distribution to be discussed here does not appear to be one of the earlier ones, although it is fair to say that in common with earlier ones, it is not the 'perfect' bivariate exponential.

The double Laplace transform of the joint p.d.f. of X_i and X_{i+1} is easily calculated using (1.1); the required expectation is

$$
\begin{aligned}
E\left\{e^{-s_1X_i-s_2X_{i+1}}\right\} &= f^{**}_{X_i,X_{i+1}}(s_1,s_2)\\
&= E\left\{e^{-\beta_2s_1\varepsilon_i-\beta_2s_2\varepsilon_{i+1}}\right\}\beta_2^2+E\left\{e^{-\beta_2s_1\varepsilon_i-s_2(\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2})}\right\}\beta_2\beta_1(1-\beta_2)\\
&+E\left\{e^{-\beta_2s_1\varepsilon_i-s_2(\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3})}\right\}\beta_2(1-\beta_2)(1-\beta_1)\\
&+E\left\{e^{-s_1(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1})-\beta_2s_2\varepsilon_{i+1}}\right\}\beta_1\beta_2(1-\beta_2)\\
&+E\left\{e^{-s_1(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1})-s_2(\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2})}\right\}\beta_1^2(1-\beta_2)^2\\
&+E\left\{e^{-s_1(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1})-s_2(\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3})}\right\}\beta_1(1-\beta_2)^2(1-\beta_1)\\
&+E\left\{e^{-s_1(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2})-\beta_2s_2\varepsilon_{i+1}}\right\}\beta_2(1-\beta_2)(1-\beta_1)\\
&+E\left\{e^{-s_1(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2})-s_2(\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2})}\right\}\beta_1(1-\beta_2)^2(1-\beta_1)\\
&+E\left\{e^{-s_1(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2})-s_2(\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3})}\right\}(1-\beta_2)^2(1-\beta_1)^2
\end{aligned}
$$

which can be written

$$
\begin{aligned}
f^{**}_{X_i,X_{i+1}}(s_1,s_2) = {} & \psi(\beta_2 s_1)\{\beta_2^2\psi(\beta_2 s_2) \\
& +\beta_1\beta_2(1-\beta_2)[\psi(\beta_2 s_2)\psi(\beta_1 s_2)+\psi(\beta_1 s_1+\beta_2 s_2)] \\
& +\beta_2(1-\beta_1)(1-\beta_2)[\psi(\beta_2 s_2)\psi(\beta_1 s_2)\psi(s_2)+\psi(\beta_1 s_1+\beta_2 s_2)\psi(s_1)] \\
& +\beta_1^2(1-\beta_2)^2[\psi(\beta_1 s_1+\beta_2 s_2)\psi(\beta_1 s_2)] \\
& +\beta_1(1-\beta_2)^2(1-\beta_1)[\psi(\beta_1 s_1+\beta_2 s_2)\psi(\beta_1 s_2)\psi(s_2)+\psi(\beta_1 s_1+\beta_2 s_2)\psi(s_1+\beta_1 s_2)] \\
& +(1-\beta_1)^2(1-\beta_2)^2[\psi(\beta_2 s_1)\psi(\beta_1 s_1+\beta_2 s_2)\psi(s_1+\beta_1 s_2)\psi(s_2)]\} \\
= {} & \frac{\lambda^2\{(\lambda+\beta_1 s_1+\beta_1 s_2)(\lambda+\beta_2 s_1+\beta_2 s_2)(\lambda+s_1)+\beta_2(1-\beta_1)s_1 s_2[\beta_2\lambda(1+\beta_1)-\beta_1\lambda-\beta_2 s_1-\beta_1(1-\beta_2)s_2]\}}{(\lambda+\beta_2 s_1)(\lambda+s_1)(\lambda+s_2)(\lambda+s_1+\beta_1 s_2)(\lambda+\beta_1 s_1+\beta_2 s_2)}
\end{aligned}
\tag{7.1}
$$

We note that (7.1) is not symmetrical in s_1 and s_2, and this is to be expected since the process is not time reversible; this is one feature which distinguishes it from earlier bivariate exponentials. The backward moving average model (1.2) corresponding to (1.1) has the joint interval distribution which is specified by (7.1) with s_1 and s_2 interchanged.

An explicit form of the joint distribution (7.1) can be obtained directly, rather than by inversion of the transform which is less informative. By the structure of the model the joint distribution of

(X_i, X_{i+1}) is a mixture of the joint distributions of

$$(\beta_2\varepsilon_i, \beta_2\varepsilon_{i+1}),\ (\beta_2\varepsilon_i, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}),\ (\beta_2\varepsilon_i, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}),$$

$$(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}, \beta_2\varepsilon_{i+1}),\ (\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}),$$

$$(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}),\ (\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2}, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}),$$

$$(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2}, \beta_2\varepsilon_{i+1}),\ (\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2}, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}).$$

with corresponding probabilities

$$\beta_2^2,\ \beta_1\beta_2(1-\beta_2),\ \beta_2(1-\beta_2)(1-\beta_1),\ \beta_1\beta_2(1-\beta_2),\ \beta_1(1-\beta_2)^2(1-\beta_1),$$

$$\beta_1^2(1-\beta_2)^2,\ \beta_1(1-\beta_2)^2(1-\beta_1),\ \beta_2(1-\beta_2)(1-\beta_1),\ (1-\beta_2)^2(1-\beta_1)^2.$$

These joint p.d.f.'s can be listed in an obvious notation as follows:

$$f_{\beta_2\varepsilon_i, \beta_2\varepsilon_{i+1}}(x,y)=(\lambda/\beta_2)^2 \exp(-\lambda x/\beta_2)\exp(-\lambda y/\beta_2) \qquad (x,y>0)$$

$$f_{\beta_2\varepsilon_i, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}}(x,y)=\lambda^2/[\beta_2(\beta_1-\beta_2)]$$
$$\cdot\exp(-\lambda x/\beta_2)\left[\exp(-\lambda y/\beta_1)-\exp(-\lambda y/\beta_2)\right] \qquad (x,y>0)$$

$$f_{\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}, \beta_2\varepsilon_{i+1}}(x,y)=(\lambda/\beta_2)^2\left\{\exp\left[-\frac{\lambda}{\beta_2}(x-\beta_1 y/\beta_2)\right]\exp(-\lambda y/\beta_2)\right\} \quad (\beta_2 x>\beta_1 y>0)$$

$$f_{\beta_2\varepsilon_i, \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}}(x,y)=\lambda^2\left[(1-\beta_2)(\beta_1-\beta_2)\right]^{-1}$$
$$\cdot\exp(-\lambda x/\beta_2)\left\{\exp(-\lambda y/\beta_2)-\exp(-\lambda y/\beta_1)-\exp(-\lambda y)+\exp\left[-\lambda y\left(1+\frac{1}{\beta_1}-\frac{1}{\beta_2}\right)\right]\right\} \quad (x,y>0)$$

The other terms are more difficult. For example, take

$$(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1},\ \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2})$$

Let $x=\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}$, $y=\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}$, $z=\beta_1\varepsilon_{i+2}$, thus

$$\varepsilon_{i+2}=z/\beta_1,\ \varepsilon_{i+1}=(y-z)/\beta_2,\ \varepsilon_i=[x-(\beta_1 y-\beta_1 z)/\beta_2]/\beta_2,$$

and the Jacobian equals to

$1/\beta_2^2\beta_1$, and $\beta_2 x=\beta_2^2\varepsilon_i+\beta_2\beta_1\varepsilon_{i+1}=\beta_2^2\varepsilon_i+\beta_1 y-\beta_1 z$ which implies that

$$\beta_2 x>\beta_1 y-\beta_1 z \Rightarrow \beta_1 z>\beta_1 y-\beta_2 x \Rightarrow z>y-\beta_2 x/\beta_1.$$

Thus when $\beta_2 x>\beta_1 y$, we integrate with respect to z from zero to y, but when $\beta_2 x<\beta_1 y$, we integrate with respect to z from $y-\beta_2 x/\beta_1$ to y. Hence we have:

when $\beta_2 x>\beta_1 y>0$

$$f_{\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1},\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}}(x,y)=\lambda^2(\beta_2^2-\beta_1\beta_2+\beta_1^2)^{-1}\cdot\exp(-\lambda x/\beta_2)\cdot$$
$$\{\exp[-\lambda y(\beta_2-\beta_1)\beta_2^{-2}]-\exp(-\lambda y/\beta_1)\}$$

when $\beta_1 y>\beta_2 x>0$

$$f_{\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1},\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}}(x,y)=\lambda^2(\beta_2^2-\beta_1\beta_2+\beta_1^2)^{-1}\cdot\exp(-\lambda y/\beta_1)\cdot$$
$$\{\exp[-\lambda x(1-\beta_2/\beta_1)/\beta_1]-\exp(-\lambda x/\beta_2)\}$$

For $(\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1},\ \beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3})$

Let $x=\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}$, $y=\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}$, $z=\beta_1\varepsilon_{i+2}$, $w=\varepsilon_{i+3}$; thus

$$\varepsilon_{i+3}=w,\ \varepsilon_{i+2}=z/\beta_1,\ \varepsilon_{i+1}=(y-z-w)/\beta_2,\ \varepsilon_i=[x-(\beta_1 y-\beta_1 z-\beta_1 w)\beta_2]/\beta_2;$$

and Jacobian $=1/\beta_2^2\beta_1$; also $\beta_2 x=\beta_2^2\varepsilon_i+\beta_1\beta_2\varepsilon_{i+1}=\beta_2^2\varepsilon_i+\beta_1 y-\beta_1 z-\beta_1 w$, which

implies that $\beta_2 x > \beta_1 y - \beta_1 z - \beta_1 w \Rightarrow \beta_1 w > \beta_1 y - \beta_1 z - \beta_1 x \Rightarrow w > y - z - \beta_2 x/\beta_1$; for $y < z + \beta_2 x/\beta_1$, integrate w from zero to y, for $y > z + \beta_2 x/\beta_1$, integrate w from $y - z - \beta_2 x/\beta_1$ to y; and in the 2nd step, since $y < z + \beta_2 x/\beta_1$ implies that $z > y - \beta_2 x/\beta_1$, if $\beta_1 y < \beta_2 x$, integrate z from zero to y, if $\beta_1 y > \beta_2 x$, integrate z from $y - \beta_2 x/\beta_1$ to y; also $y > z + \beta_2 x/\beta_1 \Rightarrow z < y - \beta_2 x/\beta_1$, thus if $\beta_2 x < \beta_1 y$, integrate z from zero to $y - \beta_2 x/\beta_1$, if $\beta_1 y < \beta_2 x$, $f(z) = 0$; hence for the expression of $f_{\beta_2 \varepsilon_i + \beta_1 \varepsilon_{i+1},\, \beta_2 \varepsilon_{i+1} + \beta_1 \varepsilon_{i+2} + \varepsilon_{i+3}}(x,y)$,

when $z > y - \beta_2 x/\beta_1$; $\beta_2 x > \beta_1 y > 0$.

$$f(x,y) = (\beta_2 \lambda)^2 \left[(\beta_2^2 + \beta_1^2 - \beta_1 \beta_2)(\beta_2^2 + \beta_1 - \beta_2)\right]^{-1} \cdot \exp(-\lambda x/\beta_2) \cdot$$
$$\left\{\exp\left[-\lambda y(1/\beta_2 - \beta_1/\beta_2^2)\right] - \exp(-\lambda y) - \exp(-\lambda y/\beta_1) + \exp\left[-\lambda y(1 + 1/\beta_1 + \beta_1/\beta_2^2 - 1/\beta_2)\right]\right\}$$

when $z > y - \beta_2 x/\beta_1$; $\beta_1 y > \beta_2 x > 0$

$$f(x,y) = (\beta_2 \lambda)^2 \left[(\beta_2^2 + \beta_1^2 - \beta_1 \beta_2)(\beta_2^2 + \beta_1 - \beta_2)\right]^{-1} \left\{\exp(-\lambda x/\beta_2) - \exp\left[-\lambda x(1 - \beta_2/\beta_1)/\beta_1\right]\right\} \cdot$$
$$\left\{\exp\left[-\lambda y(1 + 1/\beta_1 + \beta_1/\beta_2^2 - 1/\beta_2)\right] - \exp(-\lambda y/\beta_1)\right\}$$

when $z < y - \beta_2 x/\beta_1$; $\beta_1 y < \beta_2 x$

$$f(x,y) = 0$$

when $z < y - \beta_2 x/\beta_1$; $\beta_1 y > \beta_2 x > 0$

$$f(x,y) = \lambda^2 \left[(\beta_2^2 + \beta_1 - \beta_2)(1 - \beta_1)\right]^{-1}$$
$$\left\{\exp\left[-\lambda x(1 - \beta_2)/\beta_1\right] \exp(-\lambda y) - \exp\left[(-\lambda/\beta_1)(x + y - x\beta_2/\beta_1)\right]\right\}$$
$$-(\lambda \beta_2)^2 \left[(\beta_2^2 + \beta_1^2 - \beta_1 \beta_2)(\beta_2^2 + \beta_1 - \beta_2)\right]^{-1} \left\{\exp\left[-\lambda(x/\beta_2 + y)\right]\right.$$
$$\left. - \exp\left[-\lambda x(1 - \beta_2/\beta_1)/\beta_1\right] \exp\left[-\lambda y(1 + 1/\beta_1 + \beta_1/\beta_2^2 - 1/\beta_2)\right]\right\}$$

The rest can be derived in a similar way. We thus see that the joint p.d.f. of X_i, X_{i+1} will be continuous in both variables but will have different analytical expressions over the regions $\beta_2 x > \beta_1 y$ and $\beta_2 x \leq \beta_1 y$; there appears to be no compact analytical form for $f_{X_1, X_{i+1}}(x,y)$. This is unfortunate because it makes it difficult to derive maximum likelihood estimates of the parameters λ and β in the model.

Different bivariate exponentials also can be compared through their conditional properties and so we will derive these for the present contribution. Conditional p.d.f.'s are not succinct enough, and so we concentrate on conditional moments. These may be obtained from (7.1). For instance, to obtain $E(X_i | X_{i+1}=t)$, differentiate with respect to s_1 , set $s_1=0+$, multiply by -1, invert with respect to s_2 and then divide by the marginal (exponential) density of X_{i+1} . Thus

$$E(X_1|X_{1+1}=t)=\lambda^{-1}\{1+\beta_1\beta_2+\beta_1-\beta_1(1-\beta_2)\exp[-\lambda t(1-\beta_1)/\beta_1]/(\beta_1-\beta_2)$$
$$+(\beta_1-2\beta_2+\beta_1\beta_2)\exp[-\lambda t(1-\beta_2)/\beta_2]/(\beta_1-\beta_2)\}$$

Examining this regression function more closely we see that $E(X_i|X_{i+1}=t)$ is equal to λ^{-1} for $\beta_2=0$ and β_1 equals either 0 or 1; otherwise it increases exponentially from $(\beta_1\beta_2+\beta_1)\lambda^{-1}$ to the constant value $(1+\beta_1\beta_2+\beta_1)\lambda^{-1}$ as t increases. But when $\beta_2=1$ and $\beta_1=0$, $E(X_i|X_{i+1}=t)=3/\lambda$ which is the maximum value for large t.

The conditional moment $E(X_i|X_{i-1}=t)$ can be obtained similarly by interchanging s_1 and s_2 .

VIII. SOME BASIC ASPECTS OF THE EMAk MODEL

By the constructions of EMA1 and EMA2 model, we can write the general form of EMAk as:

$$
\begin{aligned}
X_i &= \beta_k \varepsilon_i, && \text{w.p. } \beta_k \\
&= \beta_k \varepsilon_i + \beta_{k-1} \varepsilon_{i+1}, && \text{w.p. } (1-\beta_k)\beta_{k-1} \\
&= \beta_k \varepsilon_i + \beta_{k-1} \varepsilon_{i+1} + \beta_{k-2} \varepsilon_{i+2}, && \text{w.p. } (1-\beta_k)(1-\beta_{k-1})\beta_{k-2} \\
&\vdots \\
&= \beta_k \varepsilon_i + \beta_{k-1} \varepsilon_{i+1} + \dots + \beta_1 \varepsilon_{i+k-1} + \beta_{i+k} && \text{w.p. } (1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_1)
\end{aligned}
\tag{8.1}
$$

Methods of mathematical induction will be used to prove some basic properties of the EMAk model.

1. The general closed form of EMAk $(k=1,2,3,\dots)$ is

$$
X_i = \sum_{j=0}^{k} \beta_{k-j} \varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(k+1-n)}, \tag{8.2}
$$

where β_0 and $I_i^{(k+1)}$ are defined to be identically 1 for all i;

$I_i^{(m)}$ is an i.i.d. sequence of Bernoulli random variables with $I_i^{(m)}=1$ w.p. $(1-\beta_m)$, 0 otherwise for all m;

i is the serial number of the ith element of the series;

k is the order of the process; j and n are indices.

Proof: When k=1

$$X_i = \beta_1 \varepsilon_i I_i^{(2)} + \beta_0 \varepsilon_{i+1} I_i^{(1)},$$

$$= \beta_1 \varepsilon_i, \qquad \text{w.p. } \beta_1$$

$$= \beta_1 \varepsilon_i + \varepsilon_{i+1}. \qquad \text{w.p. } 1-\beta_1$$

When k=2

$$X_i = \beta_2 \varepsilon_i I_i^{(3)} + \beta_1 \varepsilon_{i+1} I_i^{(3)} I_i^{(2)} + \beta_0 \varepsilon_{i+2} I_i^{(3)} I_i^{(2)} I_i^{(1)},$$

$$= \beta_2 \varepsilon_i, \qquad \text{w.p. } \beta_2$$

$$= \beta_2 \varepsilon_i + \beta_1 \varepsilon_{i+1}, \qquad \text{w.p. } (1-\beta_2)\beta_1$$

$$= \beta_2 \varepsilon_i + \beta_1 \varepsilon_{i+1} + \varepsilon_{i+2}. \qquad \text{w.p. } (1-\beta_2)(1-\beta_1)$$

Assume the result is also true when k=m then, when k=m+1

$$X_i = \sum_{j=0}^{m} \beta_{m-j} \varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(m+1-n)} + \beta_{m+1-j} \varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(m+2-n)}$$

$$= \sum_{j=0}^{m+1} \beta_{m+1-j} \varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(m+1+1-n)}.$$

This completes the proof.

2. The distribution of the intervals $\{X_i\}$ are also exponential.

Proof: $f^*_{X_i}(s)=E(e^{-sX_i})$, i.e.

$$f^*_{X_i}(s)=E(e^{-s\beta_k\varepsilon_i})\beta_k+E(e^{-s\beta_k\varepsilon_i-s\beta_{k-1}\varepsilon_{i+1}})\beta_{k-1}(1-\beta_k)+\ldots$$

$$+E(e^{-s\beta_k\varepsilon_i-s\beta_{k-1}\varepsilon_{i+1}-\ldots-s\beta_1\varepsilon_{i+k-1}-s\varepsilon_{i+k}})(1-\beta_k)(1-\beta_{k-1})\ldots(1-\beta_1)$$

$$=\frac{\lambda\beta_k}{\lambda+\beta_k s}+\frac{\lambda^2(1-\beta_k)\beta_{k-1}}{(\lambda+\beta_k s)(\lambda+\beta_{k-1}s)}+\ldots+\frac{\lambda^{k+1}(1-\beta_1)\ldots(1-\beta_k)}{(\lambda+s)(\lambda+\beta_1 s)\ldots(\lambda+\beta_k s)} \qquad (8.3)$$

When k=1,

$$f^*_{X_i}(s)=\frac{\lambda\beta_1}{\lambda+\beta_1 s}+\frac{\lambda^2(1-\beta_1)}{(\lambda+s)(\lambda+\beta_1 s)}=\frac{\lambda}{\lambda+\beta_1 s}\left[\beta_1+\frac{\lambda(1-\beta_1)}{\lambda+s}\right]=\frac{\lambda}{\lambda+s}.$$

When k=2

$$f^*_{X_i}(s)=\frac{\lambda}{\lambda+\beta_2 s}\left[\beta_2+\frac{\lambda(1-\beta_2)(\lambda+\beta_1 s)}{(\lambda+s)(\lambda+\beta_1 s)}\right]=\frac{\lambda}{\lambda+\beta_2 s}\left[\beta_2+\frac{\lambda(1-\beta_2)}{\lambda+s}\right]=\frac{\lambda}{\lambda+s}.$$

When k=m-1, the last term of (8.3) is

$$\frac{\lambda^m(1-\beta_m)(1-\beta_{m-1})\ldots(1-\beta_2)}{(\lambda+s)(\lambda+\beta_m s)(\lambda+\beta_{m-1}s)\ldots(\lambda+\beta_2 s)}.$$

Assume the result is also true when k=m-1, then, when k=m the last two terms become

$$\frac{\lambda^m(1-\beta_m)(1-\beta_{m-1})\ldots(1-\beta_2)\beta_1}{(\lambda+\beta_m s)(\lambda+\beta_{m-1}s)\ldots(\lambda+\beta_1 s)}+\frac{\lambda^{m+1}(1-\beta_m)(1-\beta_{m-1})\ldots(1-\beta_1)}{(\lambda+s)(\lambda+\beta_m s)(\lambda+\beta_{m-1}s)\ldots(\lambda+\beta_1 s)} \qquad (8.4)$$

but all the terms before these two are still the same as k=m-1, thus simplifying (8.4) gives

$$\frac{\lambda^m(1-\beta_m)(1-\beta_{m-1})\cdots(1-\beta_2)[\beta_1(\lambda+s)+\lambda(1-\beta_1)]}{(\lambda+s)(\lambda+\beta_m s)(\lambda+\beta_{m-1}s)\cdots(\lambda+\beta_2 s)(\lambda+\beta_1 s)}$$

$$=\frac{\lambda^m(1-\beta_m)(1-\beta_{m-1})\cdots(1-\beta_2)}{(\lambda+s)(\lambda+\beta_m s)(\lambda+\beta_{m-1}s)\cdots(\lambda+\beta_2 s)}.$$

which is exactly the last term of $f^*_{X_i}(s)$ when k=m-1. Hence, we proved that if the result is true when k=m-1, then the result is also true when k=m. This completes the proof.

3. The jth order serial correlation of EMAk is

$$\rho_j^{(k)}=\sum_{i=1}^{k-j+1}\beta_{k+1-i}\prod_{m=0}^{i-1}(1-\beta_{k+1-m})\beta_{k+1-j-i}\prod_{n=1}^{i+j-1}(1-\beta_{k+1-n}), \quad \text{for } 1\leq j\leq k$$

$$=0 \qquad \text{for } k<j$$

where $\beta_0=1$, and $\beta_{k+1}=0$.

Proof: By definition

$$\rho_j^{(k)}=\mathrm{corr}[X_i^{(k)},X_{i+j}^{(k)}]=\frac{\mathrm{cov}[X_i^{(k)},X_{i+j}^{(k)}]}{\{\mathrm{var}[X_i^{(k)}]\,\mathrm{var}[X_{i+j}^{(k)}]\}^{\frac{1}{2}}}$$

$$=\frac{E[X_i^{(k)}X_{i+j}^{(k)}]-E[X_i^{(k)}]\,E[X_{i+j}^{(k)}]}{\{\mathrm{var}[X_i^{(k)}]\,\mathrm{var}[X_{i+j}^{(k)}]\}^{\frac{1}{2}}},$$

where $X_i^{(k)}$'s are intervals of EMAk process, and have been proved to be marginally exponentially distributed with parameter λ . Thus

$$\left\{\operatorname{var}\left[X_i^{(k)}\right]\operatorname{var}\left[X_{i+j}^{(k)}\right]\right\}^{\frac{1}{2}}=1/\lambda^2, \text{ and}$$

$$\rho_j^{(k)}=\lambda^2\left\{E\left[X_i^{(k)}X_{i+j}^{(k)}\right]-E\left[X_i^{(k)}\right]E\left[X_{i+j}^{(k)}\right]\right\}.$$

Since $X_i^{(k)}$ and $X_{i+j}^{(k)}$ are probabilistic linear combinations of i.i.d. exponential (λ) random variables ε_i and ε_{i+j} , and

$$E(\varepsilon_i\varepsilon_{i+j})=E(\varepsilon_i)E(\varepsilon_{i+j})=1/\lambda^2,$$

the only non-zero term of $\rho_j^{(k)}$ will be the sum of

$$B\lambda^2\left[E(\varepsilon_{i+j}\varepsilon_{i+j})-E(\varepsilon_{i+j})E(\varepsilon_{i+j})\right]=B\lambda^2(2-1)\lambda^{-2}=B,$$

where B is a combination of β_i and $(1-\beta_i)$, for $i=1,2,3,\ldots$

Hence when k=1, j=1,

$$\rho_1^{(1)}=\sum_{i=1}^{1}\beta_{1+1-i}\prod_{m=0}^{i-1}(1-\beta_{1+1-m})\beta_{1+1-1-i}\prod_{n=1}^{i+1-1}(1-\beta_{1+1-n})$$

$$=\beta_1(1-\beta_2)\beta_0(1-\beta_1)=\beta_1(1-\beta_1).$$

When k=2, j=1,

$$\rho_1^{(2)}=\sum_{i=1}^{2}\beta_{3-i}\prod_{m=0}^{i-1}(1-\beta_{3-m})\beta_{2-i}\prod_{n=1}^{i}(1-\beta_{3-n})$$

$$=\beta_2(1-\beta_3)\beta_1(1-\beta_2)+\beta_1(1-\beta_3)(1-\beta_2)\beta_0(1-\beta_2)(1-\beta_1)$$

$$=\beta_2\beta_1(1-\beta_2)+\beta_1(1-\beta_2)^2(1-\beta_1)=\beta_1(1-\beta_2)-\left[\beta_1(1-\beta_2)\right]^2.$$

When k=2, j=2,

$$\rho_2^{(2)} = \sum_{i=1}^{1} \beta_{3-i} \prod_{m=0}^{i-1}(1-\beta_{3-m})\beta_{1-i} \prod_{n=1}^{i+1}(1-\beta_{3-n})$$

$$= \beta_2(1-\beta_3)\beta_0(1-\beta_2)(1-\beta_1) = \beta_2(1-\beta_2)(1-\beta_1).$$

When k=h, assume the result is also true, then, when k=h+1, j≤h+1,

$$\rho_j^{(k)} = \sum_{i=1}^{h-j+1} \beta_{h+1-i} \prod_{m=0}^{i-1}(1-\beta_{h+1-m})\beta_{h+1-j-i} \prod_{n=1}^{i+j-1}(1-\beta_{h+1-n})$$

$$+\beta_{h+1+1-(h+1-j+1)} \prod_{m=0}^{i-1}(1-\beta_{h+1+1-m})\beta_{h+1+1-j-i} \prod_{n=1}^{i+j-1}(1-\beta_{h+1+1-n})$$

$$= \sum_{i=1}^{k+1-j} \beta_{k+1-i} \prod_{m=0}^{i-1}(1-\beta_{k+1-m})\beta_{k+1-j-i} \prod_{n=1}^{i+j-1}(1-\beta_{k+1-n}).$$

This completes the proof.

4. All the correlations are non-negative and bounded above by 1/4.

Proof: From above

$$\rho_j^{(k)} = \sum_{i=1}^{k-j+1} \beta_{k+1-i} \prod_{m=0}^{i-1}(1-\beta_{k+1-m})\beta_{k+1-j-i} \prod_{n=1}^{i+j-1}(1-\beta_{k+1-n}) \qquad 1 \leq j \leq k$$

$$= \beta_k\beta_{k-j}(1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_{k+1-j})$$

$$+\beta_{k-1}(1-\beta_k)\beta_{k-j-1}(1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_{k-j})$$

$$+\beta_{k-2}(1-\beta_k)(1-\beta_{k-1})\beta_{k-j-2}(1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_{k-j-1})$$

$$\vdots$$

$$+\beta_{j+1}(1-\beta_k)\dots(1-\beta_{j+2})\beta_1(1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_2)$$

$$+\beta_j(1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_{j+1})(1-\beta_k)(1-\beta_{k-1})\dots(1-\beta_1).$$

When k=1, j=1, $\rho_1^{(1)}=\beta_1(1-\beta_1)$,

min. value=0 at $\beta_1=0$ or 1,

max. value=1/4 at $\beta_1=1/2$.

When k=2, j=1, $\rho_1^{(2)}=\beta_1(1-\beta_2)-[\beta_1(1-\beta_2)]^2$,

min. value=0 at $\beta_1=0$ or $\beta_2=1$,

max. value=1/4 at $\beta_1(1-\beta_2)=1/2$.

When k=2, j=2, $\rho_2^{(2)}=\beta_2(1-\beta_1)(1-\beta_2)$,

min. value=0 at $\beta_2=0$ or $\beta_1=1$ or $\beta_2=1$,

max. value=1/4 at $\beta_2=1/2$ and $\beta_1=0$.

When $1\leq j\leq k$, min. value=0 at $\beta_m=0$, m=j,j+1,...,k

or $\beta_m=1$, m=k,or k-1,...or k-j+1,

max. value=1/4 at $\beta_m=1/2$ and $\beta_n=0$ for $m\neq n$,

where m=k,k-1,k-2,...,j.

5. Define the i+1st element of EMAk to be

$$X_{i+1}^{(k)}=\begin{cases}\beta_k\varepsilon_{i+1}, & \text{w.p. } \beta_k\\ \beta_k\varepsilon_{i+1}+\beta_{k-1}\varepsilon_{i+2}, & \text{w.p. } (1-\beta_k)\beta_{k-1}\\ \beta_k\varepsilon_{i+1}+\beta_{k-1}\varepsilon_{i+2}+\beta_{k-2}\varepsilon_{i+3}, & \text{w.p.} (1-\beta_k)(1-\beta_{k-1})\beta_{k-2}\\ \vdots & \\ \beta_k\varepsilon_{i+1}+\beta_{k-1}\varepsilon_{i+2}+\dots+\beta_2\varepsilon_{i+k-1}+\beta_1\varepsilon_{i+k}+\varepsilon_{i+k+1}. & \text{w.p.} (1-\beta_k)(1-\beta_{k-1})\\ & \quad\dots(1-\beta_2)(1-\beta_1)\end{cases}$$

and define the ith tlement of the k+1st order process to be $X_i^{(k+1)}$,

then we can write

$$X_i^{(k+1)}=\begin{cases}\beta_{k+1}\varepsilon_i, & \text{w.p. } \beta_{k+1}\\ \beta_{k+1}\varepsilon_i+X_{i+1}^{(k)}. & \text{w.p. } 1-\beta_{k+1}\end{cases}\quad (0\leq\beta_{k+1}\leq 1;\ i=0,\pm1,\pm2,\dots)\quad (8.5)$$

Proof:

When k=1, $X_{i+1}^{(1)}=\beta_1\varepsilon_{i+1}$, w.p. β_1

$=\beta_1\varepsilon_{i+1}+\varepsilon_{i+2}$. w.p. $1-\beta_1$

Then $X_i^{(2)}=\beta_2\varepsilon_i$, w.p. β_2

$=\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}$, w.p. $(1-\beta_2)\beta_1$

$=\beta_2\varepsilon_i+\beta_1\varepsilon_{i+1}+\varepsilon_{i+2}$. w.p. $(1-\beta_2)(1-\beta_1)$

When k=2, $X_{i+1}^{(2)}=\beta_2\varepsilon_{i+1}$, w.p. β_2

$=\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}$, w.p. $(1-\beta_2)\beta_1$

$=\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}$. w.p. $(1-\beta_2)(1-\beta_1)$

Then

$$
\begin{aligned}
X_i^{(3)} &= \beta_3\varepsilon_i, && \text{w.p. } \beta_3 \\
&= \beta_3\varepsilon_i+\beta_2\varepsilon_{i+1}, && \text{w.p. } (1-\beta_3)\beta_2 \\
&= \beta_3\varepsilon_i+\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}, && \text{w.p. } (1-\beta_3)(1-\beta_2)\beta_1 \\
&= \beta_3\varepsilon_i+\beta_2\varepsilon_{i+1}+\beta_1\varepsilon_{i+2}+\varepsilon_{i+3}. && \text{w.p. } (1-\beta_3)(1-\beta_2)(1-\beta_1)
\end{aligned}
$$

When k=m-1, assume it is also true, then when k=m, do the same job, will get exact the correct result, this completes the proof.

Note that this expression is not convenient for the purpose of examining the properties of the EMAk process, since $X_i^{(k)}$ and ε_i are dependent; however, it may be used to generate the $\{X_i\}$ sequences.

IX. CONCLUSIONS

1. Both estimators of β are not very good, since the bias terms are very large when β approaches to zero or one. But $\hat{\hat{\beta}}$ looks pretty nice when β is in the interval (0.1, 0.8).
2. Estimation of β in the EMA2 process is rather difficult, because it is impossible to get the unique value of the estimators.
3. In successive stages of queueing lines, all the waiting time (waiting time in the queue plus the service time) in each stage will not be independent; this is the basic purpose of constructing this model, the size of the order k depends on the number of stages.
4. The general expression of the EMAk model is not convenient for the purpose of examining the properties of the EMAk process, since $X_i^{(k)}$ and ε_i are dependent; however, it may be used to generate the $\{X_i\}$ sequences.

APPENDIX A

METHODS OF GETTING JOINT EXPECTATIONS

1. The standard way to calculate the expectation of two or more jointly distributed random variables is to integrate the function with respect to the joint p.d.f. of the random variables; e.g.

$$E(XY)=\int_X \int_Y xy\, f_{X,Y}(x,y)\, dxdy.$$

This is not convenient for the expectations we require.

2. In the EMA1 model, a better method of getting joint expectations is to write out the expressions from the basic construction and compute them directly. For example:

$$
\begin{aligned}
X_i &= \beta\varepsilon_i && \text{w.p. } \beta, \\
&= \beta\varepsilon_i+\varepsilon_{i+1} && \text{w.p. } (1-\beta).
\end{aligned}
$$

$$
\begin{aligned}
X_{i+1}^2 &= \beta^2\varepsilon_{i+1}^2 && \text{w.p. } \beta, \\
&= \beta^2\varepsilon_{i+1}^2+2\beta\varepsilon_{i+1}\varepsilon_{i+2}+\varepsilon_{i+2}^2 && \text{w.p. } (1-\beta),
\end{aligned}
$$

thus

$$
\begin{aligned}
X_iX_{i+1}^2 &= \beta^3\varepsilon_i\varepsilon_{i+1}^2 && \text{w.p. } \beta^2, \\
&= \beta^3\varepsilon_i\varepsilon_{i+1}^2+\beta^2\varepsilon_{i+1}^3 && \text{w.p. } \beta(1-\beta), \\
&= \beta^3\varepsilon_i\varepsilon_{i+1}^2+2\beta^2\varepsilon_i\varepsilon_{i+1}\varepsilon_{i+2}+\beta\varepsilon_i\varepsilon_{i+2}^2 && \text{w.p. } \beta(1-\beta), \\
&= \beta^3\varepsilon_i\varepsilon_{i+1}^2+2\beta^2\varepsilon_i\varepsilon_{i+1}\varepsilon_{i+2}+\beta^2\varepsilon_{i+1}^3+2\beta\varepsilon_{i+1}^2\varepsilon_{i+2}+\beta\varepsilon_i\varepsilon_{i+2}^2+\varepsilon_{i+1}\varepsilon_{i+2}^2. && \text{w.p. } (1-\beta)^2
\end{aligned}
$$

By direct computation, we have:

$$
\begin{aligned}
E(X_iX_{i+1}^2X_{i+2}^2) &= 4\beta^5/\lambda^5 && \text{w.p. } \beta^3, \\
&= (12\beta^5+28\beta^4+28\beta^3)/\lambda^5 && \text{w.p. } \beta^2(1-\beta), \\
&= (12\beta^5+56\beta^4+76\beta^3+52\beta^2+4\beta)/\lambda^5 && \text{w.p. } \beta(1-\beta)^2, \\
&= (4\beta^5+28\beta^4+48\beta^3+52\beta^2+16\beta^2+4)/\lambda^5 && \text{w.p. } (1-\beta)^3.
\end{aligned}
$$

Combine the above gives

$$E(X_i X_{i+1}^2 X_{i+2}^2) = (4+4\beta+20\beta^2-20\beta^3+8\beta^5-12\beta^6)/\lambda^5.$$

All the expressions listed in APPENDIX B were computed in this way.

3. Take the derivative of the Laplace transform of the joint p.d.f. with respect to s_i, and then setting s_i equal to zero will also give the joint expectations, e.g. Lawrance and Lewis gave the general expression of Laplace transform of the joint p.d.f. of r adjacent intervals. [Ref. 1, p.17]

   Converting it gives:

$$f^{****}_{X_i X_{i+1} X_{i+2} X_{i+3}}(s_1,s_2,s_3,s_4) = \frac{\lambda^4(\lambda+\beta s_1+\beta s_2)(\lambda+\beta s_2+\beta s_3)(\lambda+\beta s_3+\beta s_4)}{(\lambda+\beta s_1)(\lambda+\beta s_2)(\lambda+\beta s_3)(\lambda+s_1+\beta s_2)(\lambda+s_2+\beta s_3)(\lambda+s_3+\beta s_4)(\lambda+s_4)}$$

   Take the derivative of this with respect to s_1 twice, s_2 once, s_3 twice, and s_4 once. Then set $s_i=0$ $(i=1,2,3,4)$ to get $E(X_i^2 X_{i+1} X_{i+2}^2 X_{i+3})$. Note that when the order of the derivative is odd, one should change the sign of the expression. This is a messy job by hand but one done easily by computer.

4. An alternative way is to use "cumulants" or "semi-invariants" [Refs. 8, p.253 and 11, p.55-93]. Let L be the Laplace transform of the joint p.d.f. and $L^*=\log L$. Let L^*_{2112} denote the derivative of L^* with respect to s_1 twice, s_2 once, s_3 once, and s_4 twice. Since $L^*_1=L_1/L$ and $L^*_2=[L_2 \cdot L-(L_1)^2]/L^2$ and $L(0)=1$, these imply that $L^*_1(0)=L_1(0)=-E(X)$ and $L^*_2(0)=L_2(0)-[L_1(0)]^2=\mathrm{var}(X)$.

If we denote $L^*_{jm}(0) = K_{jm}$ and $E(X_i^j X_{i+1}^m) = E_{jm}$ we get from this relationship the following:

$$E_{11} = K_{11} + K_1^2 ;$$

$$E_{21} = -K_{21} - K_1(2K_{11} + K_{20}) - K_1^3 ;$$

$$E_{12} = -K_{12} - K_1(2K_{11} + K_{02}) - K_1^3 ;$$

$$E_{22} = K_{22} + 2K_1(K_{12} + K_{21}) + 2K_{11}^2 + K_{20}K_{02} + 2K_1^2(2K_{11} + K_{20}) + K_1^4 ;$$

$$E_{13} = K_{13} + K_1(K_{03} + 3K_{12}) + 3K_{11}K_{02} + 3K_1^2(K_{11} + K_{02}) + K_1^4 ;$$

$$E_{31} = K_{31} + K_1(K_{30} + 3K_{21}) + 3K_{11}K_{20} + 3K_1^2(K_{11} + K_{20}) + K_1^4 ;$$

where $K_1 = K_{10} = K_{01} = -1/\lambda = -E_{01} = -E_{10}$.

Also $K_{11} = (\beta - \beta^2)/\lambda^2$,

$K_{20} = K_{02} = 1/\lambda^2$,

$K_{12} = (2\beta^3 - 2\beta)/\lambda^3$,

$K_{21} = (2\beta^3 - 2\beta^2)/\lambda^3$,

$K_{22} = (6\beta^2 - 6\beta^4)/\lambda^4$,

$K_{03} = K_{30} = -2/\lambda^3$.

APPENDIX B

LIST OF USEFUL JOINT EXPECTATIONS

$E(X_i X_{i+1}) = (1+\beta-\beta^2)/\lambda^2.$

$E(X_i^2 X_{i+1}) = (2+4\beta-2\beta^2-2\beta^3)/\lambda^3.$

$E(X_i X_{i+1}^2) = (2+2\beta-2\beta^3)/\lambda^3.$

$E(X_i^2 X_{i+1}^2) = (4+8\beta+8\beta^2-14\beta^3-2\beta^4)/\lambda^4.$

$E(X_i^3 X_{i+1}) = (6+24\beta-18\beta^2-6\beta^4)/\lambda^4.$

$E(X_i X_{i+1}^3) = (6+12\beta^2-6\beta^3-6\beta^4)/\lambda^4.$

$E(X_i^3 X_{i+1}^2) = (12+36\beta+60\beta^2-48\beta^3-36\beta^4-12\beta^5)/\lambda^5.$

$E(X_i^2 X_{i+1}^3) = (12+24\beta+24\beta^2+12\beta^3-48\beta^4-12\beta^5)/\lambda^5.$

$E(X_i^3 X_{i+1}^3) = (36+108\beta+180\beta^2+216\beta^3-324\beta^4-144\beta^5-36\beta^6)/\lambda^6.$

$E(X_i^4 X_{i+1}) = (24+96\beta-24\beta^2-24\beta^3-48\beta^4+24\beta^5-24\beta^6)/\lambda^5.$

$E(X_i X_{i+1}^4) = (24+24\beta-24\beta^3+48\beta^4-48\beta^5)/\lambda^5.$

$E(X_i^5 X_{i+1}) = (120+600\beta-120\beta^2-120\beta^3-120\beta^4-120\beta^5-120\beta^6)/\lambda^6.$

$E(X_i X_{i+1}^5) = (120+120\beta-120\beta^6)/\lambda^6.$

$E(X_i^2 X_{i+1}^4) = (48+96\beta+96\beta^2+48\beta^3+48\beta^4-240\beta^5-48\beta^6)/\lambda^6.$

$E(X_i^4 X_{i+1}^2) = (48+192\beta+432\beta^2-240\beta^3-192\beta^4-144\beta^5-48\beta^6)/\lambda^6.$

$$E(X_i^4 X_{i+1}^4) = (576+2304\beta+5184\beta^2+8640\beta^3+12096\beta^4-16704\beta^5-8064\beta^6$$
$$-2880\beta^7-576\beta^8)/\lambda^8.$$

$$E(X_i X_{i+1} X_{i+2}) = (1+\beta-\beta^3)/\lambda^3.$$

$$E(X_i^2 X_{i+1} X_{i+2}) = (2+6\beta-4\beta^2-2\beta^3)/\lambda^4.$$

$$E(X_i X_{i+1}^2 X_{i+2}) = (2+6\beta-8\beta^3+2\beta^4)/\lambda^4.$$

$$E(X_i X_{i+1} X_{i+2}^2) = (2+4\beta-2\beta^2-2\beta^3)/\lambda^3.$$

$$E(X_i^2 X_{i+1}^2 X_{i+2}) = (4+16\beta+36\beta^2-96\beta^3+96\beta^4-88\beta^5+36\beta^6)/\lambda^5.$$

$$E(X_i^2 X_{i+1} X_{i+2}^2) = (4+12\beta-4\beta^2-14\beta^3+12\beta^4-6\beta^5)/\lambda^5.$$

$$E(X_i X_{i+1}^2 X_{i+2}^2) = (4+4\beta+20\beta^2-20\beta^3+8\beta^5-12\beta^6)/\lambda^5.$$

$$E(X_i^2 X_{i+1}^2 X_{i+2}^2) = (8+32\beta+48\beta^2-56\beta^3-40\beta^4+8\beta^5+8\beta^6)/\lambda^6.$$

$$E(X_i X_{i+1}^3 X_{i+2}) = (6+6\beta+42\beta^2-54\beta^3+16\beta^5-10\beta^6)/\lambda^5.$$

$$E(X_i^2 X_{i+1}^3 X_{i+2}) = (12+96\beta-36\beta^2+60\beta^3-156\beta^4+204\beta^5-168\beta^6)/\lambda^6.$$

$$E(X_i X_{i+1}^3 X_{i+2}^2) = (12+168\beta-276\beta^2+348\beta^3-312\beta^4+72\beta^5)/\lambda^6.$$

$$E(X_i X_{i+1}^2 X_{i+2}^2 X_{i+3}) = (4+16\beta+68\beta^2-132\beta^3+104\beta^4-160\beta^5+124\beta^6+4\beta^7$$
$$-24\beta^8)/\lambda^6.$$

$$E(X_i^2 X_{i+1} X_{i+2}^2 X_{i+3}) = (4+20\beta+12\beta^2-28\beta^3-24\beta^4-4\beta^5+56\beta^6-44\beta^7$$
$$+12\beta^8)/\lambda^6.$$

$$E(X_i^2 X_{i+1} X_{i+2} X_{i+3}^2) = (4+24\beta-44\beta^2+80\beta^3-108\beta^4+60\beta^5-20\beta^6+12\beta^7$$
$$-4\beta^8)/\lambda^6.$$

$$E(X_i X_{i+1}^2 X_{i+2} X_{i+3}^2) = (4+16\beta+4\beta^2-36\beta^3+52\beta^4-64\beta^5+36\beta^6-16\beta^7$$
$$+8\beta^8)/\lambda^6.$$

$$E(X_i X_{i+1}^2 X_{i+2} X_{i+3}) = (2+8\beta-4\beta^2+2\beta^3-14\beta^4+10\beta^5-2\beta^6)/\lambda^5.$$

$$E(X_i X_{i+1} X_{i+2}^2 X_{i+3}) = (2+8\beta+2\beta^2-26\beta^3+26\beta^4-6\beta^5-12\beta^7+8\beta^8)/\lambda^5.$$

$$E(X_i X_{i+1} X_{i+2} X_{i+3}^2) = (2+6\beta-2\beta^2-4\beta^3-2\beta^4-2\beta^5)/\lambda^5.$$

$$E(X_i X_{i+1}^4 X_{i+2}) = (24+120\beta+48\beta^2-96\beta^3-168\beta^5+96\beta^6)/\lambda^6.$$

LIST OF REFERENCES

1. Lawrance, A. J. and Lewis, P. A. W. (1975). A Moving Average Exponential Point Process (EMA1). Naval Postgraduate School Report NPS55Lw75061.

2. Lawrance, A. J. (1972). Some Models for Stationary Series of Univariate Events. In Stochastic Point Processes (P. A. W. Lewis, ed.) Wiley, New York, 199-256.

3. Cox, D. R. and Lewis, P. A. W. (1966). The Statistical Analysis of Series of Events. Methuen, London and Wiley, New York.

4. Anderson, T. W. (1971). The Statistical Analysis of Time Series. Wiley, New York.

5. Box, G. E. P. and Jenkins, G. M. (1970). Time Series Analysis Forecasting and Control. Holden-Day, San Francisco.

6. Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.

7. Rosenblatt, M. (1962). Random Processes. Oxford, New York.

8. Gnedenko, B. V. (1962). The Theory of Probability. Chelsea, New York.

9. Rosenblatt, M. (1971). Markov Processes. Structure and Asymptotic Behavior. Springer-Verlag, New York.

10. Feller, W. (1966). An Introduction to Probability Theory and its Application. Volume 2. Wiley, New York.

11. Kendall, M. G. and Stuart, A. (1952). The Advanced Theory of Statistics. Volume 1. Hafner, New York.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
4.	Professor Peter A. W. Lewis, Code 55Lw Department of Operations Research Naval Postgraduate School Monterey, California 93940	3
5.	Professor Donald P. Gaver, Code 55Gv Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
6.	Assoc. Professor Paul R. Milch, Code 55Mh Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
7.	Assoc. Professor Francis R. Richards, Code 55Rh Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
8.	Assoc. Professor Donald R. Barr, Code 55Bn Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
9.	Dr. B. J. McDonald, Program Director Probability and Statistics Program Office of Naval Research Arlington, Virginia 22217	1
10.	Dr. Merlin G. Bell, (LCDR, USN) Class No. 59 Armed Forces Staff College Norfolk, Virginia 23511	1
11.	CDR. Lo, Tzy-dah Jathro, Chinese Navy, SMC 2540 Department of Operations Research Naval Postgraduate School Monterey, California 93940	1

Thesis 165189
L775 Lo
c.1
An investigation of
the properties of the
exponential moving
average point process.